

2015

Bioprospecting For Genes That Confer Biofuel Tolerance To Escherichia Coli Using A Genomic Library Approach

Timothy Tomko
University of Vermont

Follow this and additional works at: <http://scholarworks.uvm.edu/graddis>



Part of the [Biology Commons](#), and the [Power and Energy Commons](#)

Recommended Citation

Tomko, Timothy, "Bioprospecting For Genes That Confer Biofuel Tolerance To Escherichia Coli Using A Genomic Library Approach" (2015). *Graduate College Dissertations and Theses*. Paper 487.

This Thesis is brought to you for free and open access by the Dissertations and Theses at ScholarWorks @ UVM. It has been accepted for inclusion in Graduate College Dissertations and Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact donna.omalley@uvm.edu.

BIOPROSPECTING FOR GENES THAT CONFER BIOFUEL TOLERANCE TO
ESCHERICHIA COLI USING A GENOMIC LIBRARY APPROACH

A Thesis Presented

by

Timothy Tomko

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Master of Science
Specializing in Mechanical Engineering

October, 2015

Defense Date: July 14, 2015
Thesis Examination Committee:

Mary J. Dunlop, Ph.D., Advisor
Matthew Wargo, Ph.D., Chairperson
Jason Bates, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

ABSTRACT

Microorganisms are capable of producing advanced biofuels that can be used as ‘drop-in’ alternatives to conventional liquid fuels. However, vital physiological processes and membrane properties are often disrupted by the presence of biofuel and limit the production yields. In order to make microbial biofuels a competitive fuel source, finding mechanisms of improving resistance to the toxic effects of biofuel production is vital. This investigation aims to identify resistance mechanisms from microorganisms that have evolved to withstand hydrocarbon-rich environments, such as those that thrive near natural oil seeps and in oil-polluted waters.

In this study, screened the genomes of two types of bacteria, *Pseudomonas aeruginosa* and *Marinobacter aquaeolei*, looking for genes that impart biofuel tolerance when expressed in *Escherichia coli*. Both of these microbes have adapted in their respective natural environments to contain mechanisms for dealing with environmental stress. For initial work, *P. aeruginosa* was used to test our experimental design and procedure, and we validated our methods by identifying a gene, *ohr* from *P. aeruginosa*, that increased tolerance to the bio-jet fuel precursor limonene in *Escherichia coli*.

Using genomic DNA from *M. aquaeolei*, we constructed a transgenic library that we expressed in *E. coli*. We exposed cells to inhibitory levels of pinene, a monoterpene that can serve as a jet fuel precursor with chemical properties similar to existing tactical fuels. Using a sequential strategy of a fosmid library followed by a plasmid library, we were able to isolate a region of DNA from the *M. aquaeolei* genome that conferred pinene tolerance when expressed in *E. coli*. We determined that a single gene, *yceI*, was responsible for the tolerance improvements. Overexpression of this gene placed no additional burden on the host. We also tested tolerance to other monoterpenes and showed that *yceI* selectively improves tolerance.

CITATIONS

Material from this thesis has been submitted for review on June 8th, 2015 in the following form:

Tomko, T. and Dunlop M.J.. Engineering improved bio-jet fuel tolerance in *Escherichia coli* using a transgenic library from the hydrocarbon-degrader *Marinobacter aquaeolei*.

ACKNOWLEDGEMENTS

I would like to acknowledge the many faculty members, staff and fellow students who helped make my thesis possible. In particular I would like to thank Dr. Mary Dunlop for her guidance, support, and creative solutions as well as for helping me develop as a research scientist and engineer. I would also like to thank my committee members, Dr. Jason Bates and Dr. Matthew Wargo. I also want to extend my appreciation to the members of the Dunlop Lab group for their support and advice.

TABLE OF CONTENTS

	Page
CITATIONS.....	.ii
ACKNOWLEDGEMENTS.....	.iii
LIST OF FIGURESv
LIST OF TABLES.....	.vi
CHAPTER 1: INTRODUCTION	1
1.1. Microbes that Thrive in Harsh Environments	1
1.2. Bio-jet Fuels as Alternatives to Conventional Fuels	2
1.3. Challenges with and Solutions to Increased Biofuel Production.....	5
CHAPTER 2: INITIAL TESTS RELATED TO SCREENING A GENOME FOR BIOFUEL TOLERANCE.....	8
2.1. Preliminary Studies Using a <i>Pseudomonas aeruginosa</i> Library	8
2.2. Mathematical Modeling to Guide the Competition Experiment Design	9
2.3. Performing the Competition Experiment with the <i>P. aeruginosa</i> Library	12
2.3. Ohr from <i>P. aeruginosa</i> Increase Limonene Tolerance in <i>E. coli</i>	14
CHAPTER 3: FOSMID LIBRARY DESIGN AND CONSTRUCTION	16
3.1. Designing a Suitable Fosmid Library for Screening	16
3.2. Testing the <i>M. aquaeolei</i> Fosmid Library in Pinene	19

3.3. Comparing the competition and the modeling results	22
3.4. Creating and Testing a Sub-library from pCC1FOS-96	24
3.5. YceI Improves Pinene Tolerance in <i>E. coli</i>	25
CHAPTER 4: FUTURE DIRECTIONS	32
4.1. Further Work to be Considered	32
CHAPTER 5: METHODS	34
5.1. Fosmid Library Construction	34
5.2. Selection Procedure: Fosmid Library	35
5.3. Plasmid Library Construction	36
5.4. Selection Procedure: Plasmid Library	37
5.5. Subcloning Genes from the Converged Plasmid	37
5.6. Tolerance Testing Procedure	38

LIST OF FIGURES

Figure	Page
Figure 1: An illustration of catagenesis	2
Figure 2: An illustration of a distillation tower	3
Figure 3: Competition experiment design	9
Figure 4: Lotka-Volterra three strain model	10-11
Figure 5: Lotka-Volterra 100 strain model with dilutions	12
Figure 6: Identification of the <i>ohr</i> gene	14
Figure 7: Testing the tolerance benefit of expressing <i>ohr</i>	15
Figure 8: Fosmid library creation procedure	17
Figure 9: Comparison of genomic DNA samples on a gel	18
Figure 10: Initial tolerance testing of fosmid containing cells	19
Figure 11: Genomic library approach for identifying biofuel tolerance genes	11
Figure 12: Improved version of the Lotka-Volterra model	23
Figure 13: Isolated genes from the fosmid library competition	25
Figure 14: Identification of <i>yceI</i> as the pinene tolerance improving gene	26
Figure 15: Tolerance testing of <i>yceI</i> in pinene	28
Figure 16: Toxicity to cells when expressing <i>yceI</i>	29
Figure 17: Tolerance testing of <i>yceI</i> in other monoterpene fuels	30
Figure 18: Chemical structures of pinene, limonene, terpinene, and terpinolene	31

LIST OF TABLES

Table	Page
Table 1: DNA fragments recovered from <i>M. aquaeolei</i>	21

CHAPTER 1: INTRODUCTION

1.1. Microbes that Thrive in Harsh Environments

Microorganisms are capable of producing advanced biofuels that can be used as ‘drop-in’ alternatives to conventional petroleum-based liquid fuels (Fischer, Klein-Marcuschamer et al. 2008, Fortman, Chhabra et al. 2008, Lee, Chou et al. 2008). However, many of these fuels are toxic to cells, introducing an undesirable trade-off between cell survival and biofuel production. There are many natural examples of microorganisms that survive in hydrocarbon-rich environments (Rojas, Duque et al. 2001, Camilli, Reddy et al. 2010). We asked whether the genomes of these microbes could be a source of biofuel tolerance mechanisms when expressed heterologously in a biofuel production host. Microbes that tolerate hydrocarbons have been isolated near natural oil seepages and around oil spills. For example, a hydrocarbon degrading microbe, *Marinobacter aquaeolei* VT8, was isolated at the head of an offshore oil well in Vietnam (Huu, Denner et al. 1999). A previous study showed that this organism harbors two efflux pumps that can serve to improve biofuel tolerance (Dunlop, Dossani et al. 2011). That study also identified an efflux pump from *Alcanivorax borkumensis*, another hydrocarbon-degrader that thrives in oil-polluted waters. In addition to the ability to metabolize hydrocarbons, *A. borkumensis* possess multiple tolerance mechanisms, including the production of biosurfactants, efflux pumps, and niche-specific stress responses (Schneiker, dos Santos et al. 2006). Studies of the microbial communities in the Gulf of Mexico after the Deepwater Horizon oil spill showed a high proportion of γ -Proteobacteria harboring hydrocarbon-degrading genes (Hazen,

Dubinsky et al. 2010). Given the abundance of naturally hydrocarbon-tolerant microorganisms, we hypothesized that the genomes of these organisms may serve as an untapped reservoir of tolerance genes.

1.2. Bio-jet Fuels as Alternatives to Conventional Fuels

Conventional hydrocarbon based fuels are formed through the process known as catagenesis over the course of thousands and thousands of years (Schobert 2013). Through this process, shown in Figure 1, hydrogen is either lost or gained to form molecules with varying chemical properties.

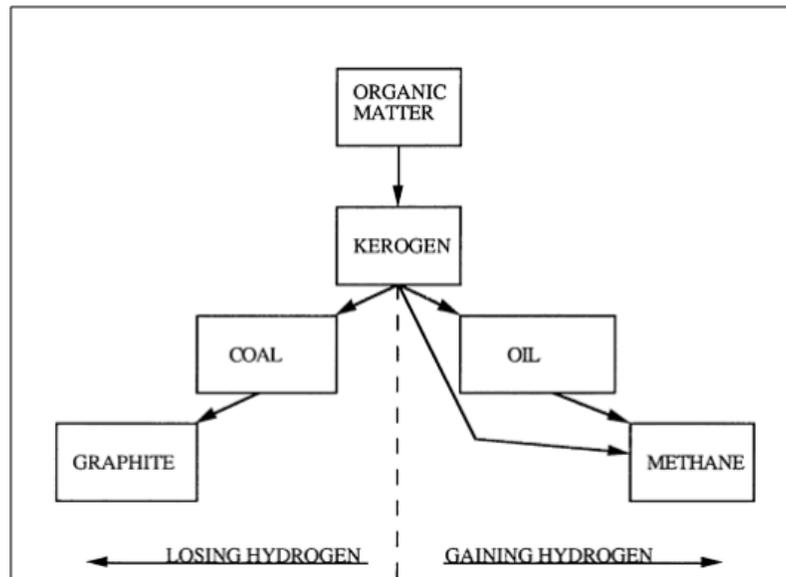


Figure 1. Shows an illustration of catagenesis, the process by which hydrocarbons gain/lose hydrogen (Schobert 2013).

An increasing hydrogen to carbon (H:C) ratio means that the molecules are becoming saturated with hydrogen and will continue this process until reaching the final state of methane, CH₄. Conversely, a decreasing H:C ratio means that hydrocarbon is becoming more dense until reaching the final state of pure carbon which is known as graphite. Since the process of catagenesis takes place over a very long time period, intermediates with varying levels of H:C ratios are found. These intermediates include hydrocarbons like coal, oil, wax, etc.

Crude oil is the starting material for many types of liquid fuels, however it must be refined via distillation before it can be used (Schobert 2013). The process of distillation (Figure 2) separates the crude oil into several “cuts” based on their boiling points.

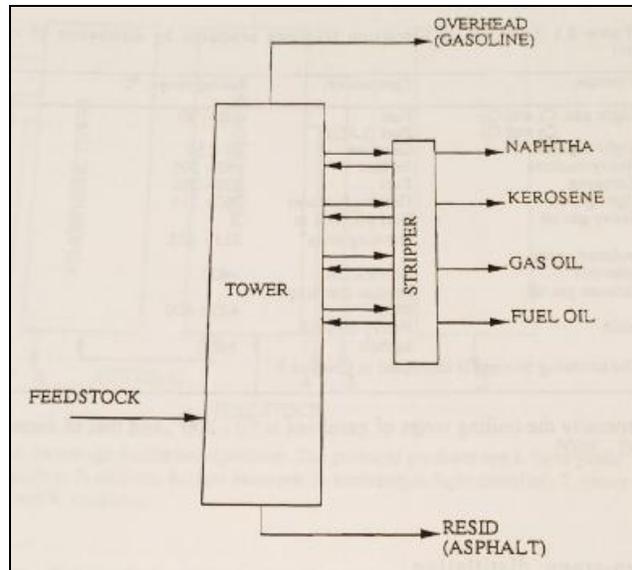


Figure 2. Shows an illustration of a distillation tower separating hydrocarbons into a variety of “cuts” based on their boiling points (Schobert 2013).

Higher boiling point hydrocarbons like resid and fuel oil have very long and complex hydrocarbon chains. Gas oil (diesel fuel) and kerosene (jet fuel) have slightly lower boiling points and therefore come out of the distillation column above the fuel oil and resid. Finally, the lightest products, naphtha and overhead gasoline are all used downstream to make conventional gasoline products.

In this study we focused specifically on improving tolerance to bio-jet fuels. Gasoline and diesel fuel can be supplemented with the use of battery power in terrestrial vehicles, however this is not a feasible option for aircraft that have much more severe weight and size limitations. Aviation fuel must have a low enough freezing point that the fuel does not gel at low temperatures associated with typical flight altitudes. Also, the fuel must burn cleanly so as to not produce large amounts of soot that could potentially damage the turbine (Schobert 2013). These considerations make

bio-jet fuels a promising, but challenging, class of biofuels to produce. Monoterpenes ($C_{10}H_{16}$) such as pinene, limonene, terpinene, and terpinolene are composed of two isoprene units (Schobert 2013) and have been shown to be excellent candidates for replacements to commercial Jet-A/A-1 fuels (Harvey, Wright et al. 2010, Brennan, Turner et al. 2012, Chuck and Donnelly 2014). Importantly, several of these bio-jet fuels have been produced by engineered microbial hosts. For example, in 2014, Sarria, *et al.* reported a pinene production pathway in *E. coli* (Sarria, Wong et al. 2014). Limonene production pathways have also been engineered in *E. coli* (Dunlop, Dossani et al. 2011, Alonso-Gutierrez, Chan et al. 2013).

1.3. Challenges with and Solutions to Increased Biofuel Production

A major challenge in the production of advanced biofuels is that toxicity limits the concentration of biofuel a cell can withstand. Vital physiological processes and membrane properties are often disrupted by the presence of these hydrocarbons; for reviews on mechanisms of biofuel toxicity and engineering approaches to mitigating toxicity see (Ramos, Duque et al. 2002, Nicolaou, Gaida et al. 2010, Dunlop 2011, Peabody, Winkler et al. 2014). Previous tolerance engineering efforts have identified several ways that biofuel tolerance can be increased. For instance, export pumps, including efflux pumps and ABC transporters, can be used to improve biofuel tolerance and export (Fischer, Klein-Marcuschamer et al. 2008, Dunlop, Dossani et al. 2011,

Doshi, Nguyen et al. 2013, Foo and Leong 2013). Additionally, several studies have shown that heat shock proteins can play an important role in improving solvent tolerance; recent examples include (Fiocco, Capozzi et al. 2007, Alsaker, Paredes et al. 2010, Reyes, Almario et al. 2011). General stress response proteins can also mitigate biofuel toxicity and have appeared in screens for tolerance genes (Weber, Polen et al. 2005, Nicolaou, Gaida et al. 2010)

Genomic library approaches have been successful at selecting for genes that improve tolerance. Typically, genomic DNA is either digested or sheared into fragment sizes appropriate for subcloning, inserted into the vector of choice, and transformed into the host. Cells containing members of the library are then subjected to a stressor and those that survive are screened using microarrays or sequencing. Most experiments related to biofuel tolerance that have used this approach use autologous libraries (for example, using genomic DNA from *E. coli* to screen in *E. coli*). Woodruff *et al.* used a multi-Scalar Analysis of Library Enrichments (SCALEs) approach to identify nine novel target genes to improve ethanol tolerance in *E. coli* (Woodruff, Pandhal et al. 2013). Other studies involving the overexpression of endogenous genes have shown improvements in isobutanol and ethanol tolerance in *Saccharomyces cerevisiae* (Hong, Lee et al. 2010), and ethanol (Nicolaou, Gaida et al. 2012) and n-butanol (Reyes, Almario et al. 2011) tolerance in *E. coli*. A small number of studies have used transgenic DNA libraries to improve tolerance in processes related to biofuel production. Zingaro *et al.* identified several tolerance mechanisms from *Lactobacillus plantarum*, that when expressed in *E. coli* show improved survival and growth under

ethanol stress (Zingaro, Nicolaou et al. 2014). Ruegg *et al.* increased ionic liquid tolerance in an *E. coli* host by screening a fosmid library from *Enterobacter lignolyticus* for tolerance genes (Ruegg, Kim et al. 2014). Genomic libraries represent an efficient way of screening many genes for desirable traits and transgenic libraries extend this approach to a potentially rich resource of tolerance genes.

There are many examples of studies where screening for genes that improve tolerance to exogenous addition of biofuel has proved successful in identifying tolerance genes that also improve biofuel yields. For example, in a recent study of isopentanol tolerance in *E. coli*, six out of eight of the genes that were found to improve tolerance to exogenous biofuel also increased production titers (Foo, Jensen et al. 2014). An efflux pump that was identified to improve tolerance to limonene also increased yield in a production strain (Dunlop, Dossani et al. 2011). A study performed in *Saccharomyces cerevisiae* showed that improved ethanol tolerance translates to a more efficient glucose to ethanol conversion rate (Alper, Moxley et al. 2006). However, there are counterexamples where improved tolerance does not lead to an increase in biofuel production, for example Atsumi *et al.* identified mutations that improved isobutanol-tolerance in *E. coli*. However their mutants did not yield higher titers of the biofuel product (Atsumi, Wu et al. 2010). Due to the burden biofuel production pathways place on the cell, it is often more straightforward to screen for tolerance improvements with exogenous biofuel addition and later incorporate the tolerance strategy into a production strain. This approach has proved successful in many previous

studies, and in this work we screen for monoterpene tolerance using exogenous addition of the biofuel.

This thesis aims to identify resistance mechanisms from microorganisms that have evolved to withstand extreme environmental conditions that can be used in engineered biofuel production strains. The goal of this research is to identify novel tolerance genes that make it possible for *E. coli* to survive in high concentrations of bio-jet fuels.

CHAPTER 2: INITIAL TESTS RELATED TO SCREENING A GENOME FOR BIOFUEL TOLERANCE

2.1. Preliminary Studies Using a *Pseudomonas aeruginosa* Library

Using a transgenic library approach, we screened for genes from microbes that have evolved to survive harsh conditions in their environment. Genes that allow these microbes to survive in the presence of harmful chemicals could be useful in increasing the tolerance of *E. coli* to biofuel. Initially, we studied *Pseudomonas aeruginosa* due to its prominence as a pathogen and its use of efflux pumps to survive in the presence of a variety of solvents (Ramos, Duque et al. 2002). *P. aeruginosa* strains expressing wild-type levels of their multidrug efflux system, MexAB-OprM, showed increased tolerance to *n*-hexane and *p*-xylene versus the strain with those genes deleted (Li, Zhang et al. 1998). We used a *P. aeruginosa* plasmid library provided by Dr. Wargo at The University of Vermont for our initial tolerance screening experiments (Wargo and Hogan 2009). Fragments of the genomic DNA of *P. aeruginosa* on plasmids is contained in the library. We transformed this library into *E. coli* MG1655 for our initial biofuel tolerance tests.

Under biofuel stress, cells compete for resource. Any cells that contain transgenic DNA that helps them to survive will out compete the other cells in the population. Over the course of several serial dilutions into fresh media and biofuel, the population should begin to converge to cells containing the same transgenic DNA, shown in Figure 3. By analyzing the sequence information from the plasmids in these cells, it is possible to identify which gene(s) the outperforming cells contain. To do the

tolerance screening, we designed a competition experiment to stress the library with biofuel added externally.

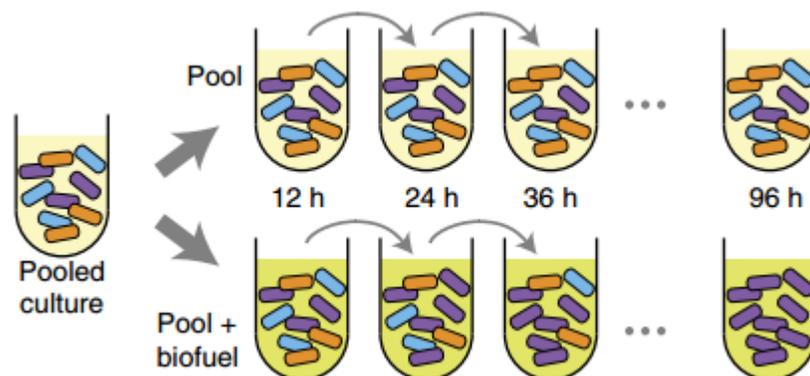


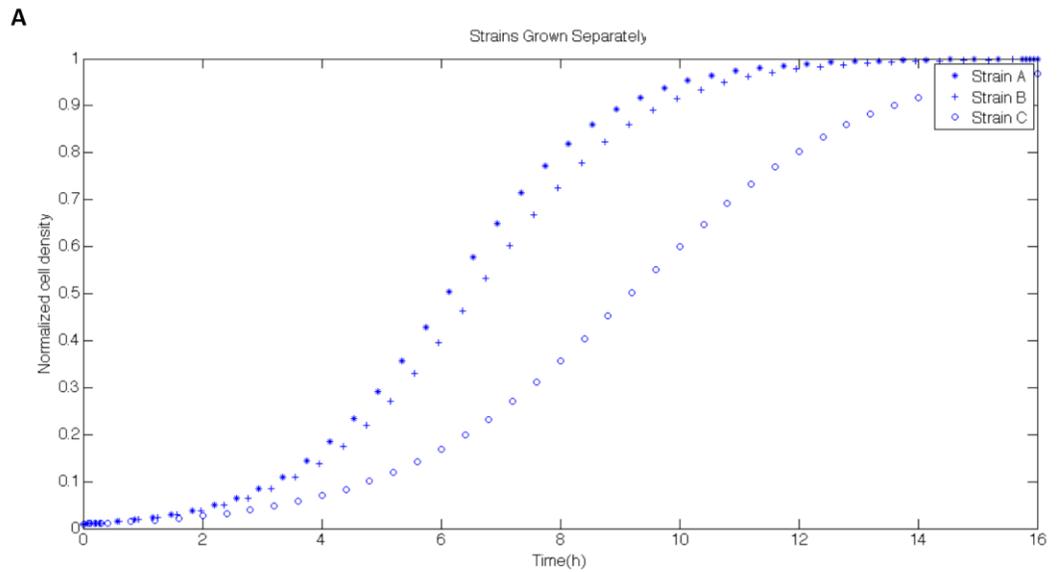
Figure 3: Competition experimental design (Dunlop 2011b). The library is stressed with biofuel over the course of 96 hours with serial dilutions into fresh media and biofuel happening every 12 hours. A control without biofuel in the media is also tested alongside the biofuel stressed culture.

2.2. Mathematical Modeling to Guide the Competition Experiment Design

Mathematical modeling using the competitive Lotka-Volterra equation was used to verify the concept of the competition experiment and the modeling work was derived from previous studies performed by Dunlop et al. (Strogatz 2005) (Dunlop, Dossani et al. 2011).

$$\frac{dN_i}{dt} = d_i N_i \left(1 - \sum_{j=1}^C \frac{d_j N_j}{d_i K} \right)$$

In the equation shown above, for the i^{th} strain, N_i is the cell density, d_i is the growth rate and K is the carrying capacity of the culture. Figure 4A shows a scenario where three different strains of bacteria with varying growth rates are grown independently of each other. In this scenario, there is no competition between the strains and each strain reaches its maximum cell density at a different time. In Figure 4B the three strains compete against each other in the same culture. The strains with the faster growth rates outperform the slower strains ($A > B > C$).



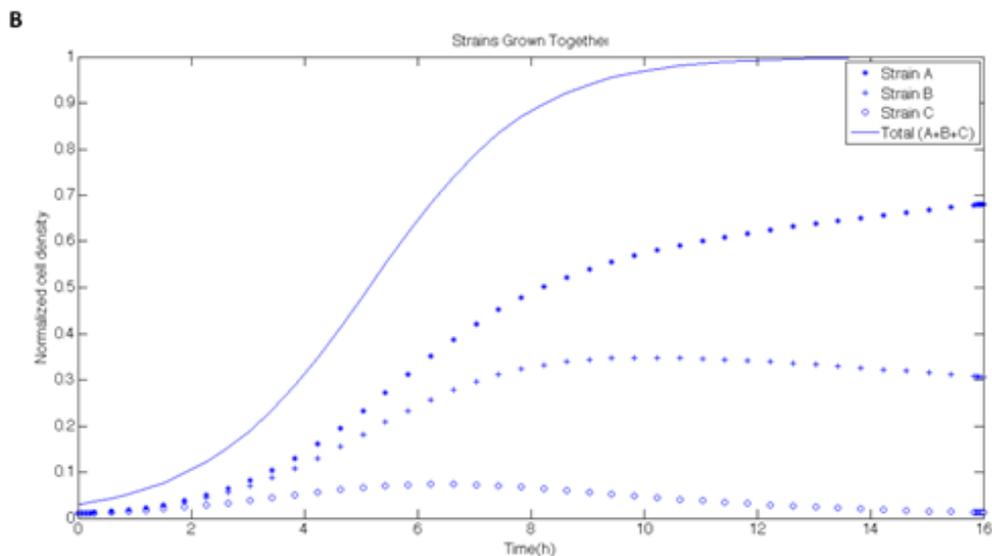


Figure 4. Model using the competitive Lotka-Volterra equation (A) Shows the scenario when there is no competition between the strains, with each strain growing independently. (B) Growing the strains in the same culture provides competition and variation in the final cell densities.

From this idea, applying a stress (such as adding biofuel) to cells expressing our library, should result in faster growing cells outcompeting the others. The model was also used to determine roughly how long to run the competition for before convergence to a single genotype would occur. Shown in Figure 5 is a model where $C = 100$ strains of bacteria and d_i was randomly distributed with a mean and standard deviation of $\mu = 0.75$ and $\sigma = 0.02$. The carrying capacity was also held constant at $K=1$. Over the course of 96 hours, the spread between the individual cell densities becomes large. This means that the population of cells is converging to the fastest growing variants.

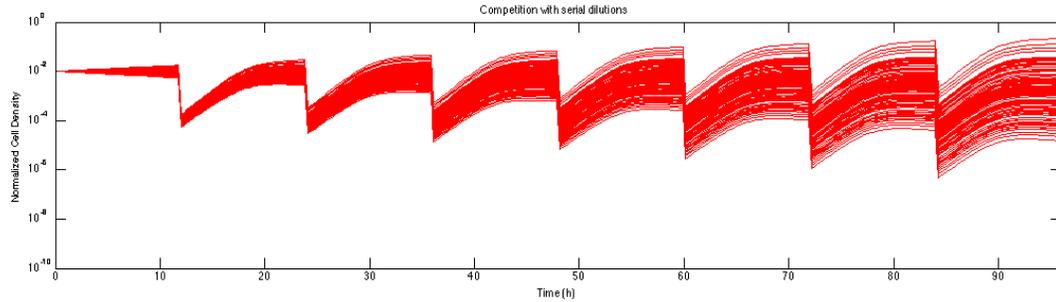


Figure 5. Model using the competitive Lotka-Volterra equation to compete 100 strains of bacteria with varying growth rates over the course of 96 hours. 1:100 serial dilutions occur every 12 hours.

Based on the results of the modeling, we decided to run the competition with our library for 96 hours, with 1:100 serial dilutions happening every 12 hours.

2.3. Performing the Competition Experiment with the *P. aeruginosa* Library

We did an initial test of the competition experiment with the *P. aeruginosa* plasmid library. The plasmid library was transformed into MG1655 *E. coli* cells and the resulting colonies were then pooled together to form the starting population of cells that would make up the library. During the experimental setup phase, it was important to try and prevent any cell growth within the population prior to the addition of biofuel. This is because tolerance mechanisms could be costly and detrimental to the cell in the absence of biofuel. These cells could very quickly be outcompeted by others in the population if biofuel was not present. Therefore, we transformed the plasmid library

into *E. coli* and used colonies directly from the agar plates for our experiments. This prevented any unwanted competition between cells before we started our testing.

We ran the competition experiment in the presence of 0.2% limonene (v/v) for 96 hours with serial dilutions into fresh media every 12 hours. This concentration of limonene was selected because it inhibited growth without completely killing *E. coli* cells not containing the library. Along with the population grown in the presence of biofuel, a separate population was grown in the absence of biofuel as a control to represent the natural competition occurring within the population. Cultures were plated every 24 hours and at the end of the competition, plasmids from five colonies on the 96 hour plate were isolated and sequenced to determine what *P. aeruginosa* DNA they contained. Interestingly, all of the sequenced plasmids from the 96 hour biofuel exposed plate contained the same five genes from *P. aeruginosa*. Each potential tolerance gene was then subcloned into a separate vector in order to determine which of the genes was responsible for the increased limonene tolerance observed in *E. coli*. We used a vector (pBbA5k) from the BioBricks series vectors, placing the gene downstream of a LacI-repressed promoter (Lee, Krupa et al. 2011). We supplemented the medium with 100 μ M of IPTG to induce expression. We did tolerance testing on each gene and, as shown in Figure 6, gene 4 was much better than the rest at improving *E. coli*'s tolerance to 0.2% limonene. The negative control (MG1655/pBbA5k-rfp) contained red fluorescent protein (*rfp*) instead of transgenic DNA. It showed poor growth when exposed to limonene. The positive control, the winning plasmid obtained after 96 hours of the competition experiment was retransformed into fresh MG1655

cells to eliminate the possibility that any observed tolerance increases were because of mutations on the genome. The positive control plasmid containing genes 1-5 grew well in the presence of biofuel and showed similar tolerance levels to the cells expressing only gene 4.

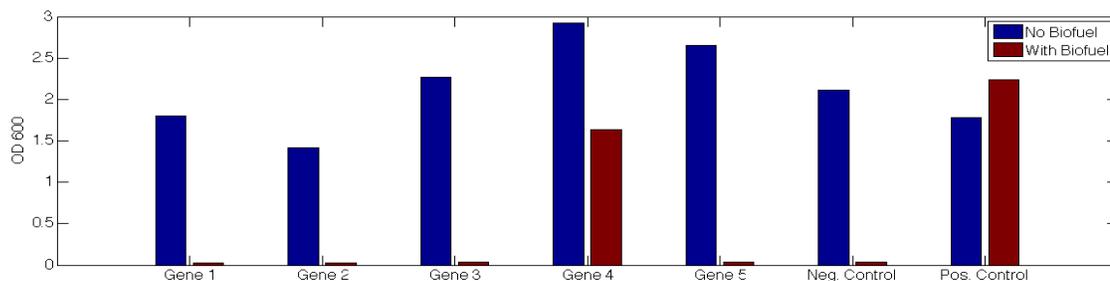


Figure 6: Tolerance testing in 0.2% limonene of individual genes recovered from the *P. aeruginosa* library after the competition experiment. Limonene was added at time zero and gene expression was induced with 100 μ M of IPTG at the same time. OD600 measurements were taken after 16 hours of growth. The negative control (pBbA5k-rfp) contains only a red florescent protein gene. The positive control contains all five genes recovered from the competition experiment.

2.4. Ohr from *P. aeruginosa* Increase Limonene Tolerance in *E. coli*

Gene 4 encodes an organic hydroperoxide resistance protein Ohr (YP_790323.1), and was clearly the gene responsible for the improved limonene tolerance. In *P. aeruginosa*, Ohr works to neutralize oxidative species produced during bacterial aerobic respiration or by the host immune system as a defense mechanism

(Lesniak, Barton et al. 2002). It may function by converting the oxidative molecule into a less toxic form. In this case, Ohr may alleviate oxidative stress in *E. coli* caused by the addition of limonene. Further testing of *ohr* in Figure 7 showed improved growth versus the *rfp* negative control up to 0.4% limonene.

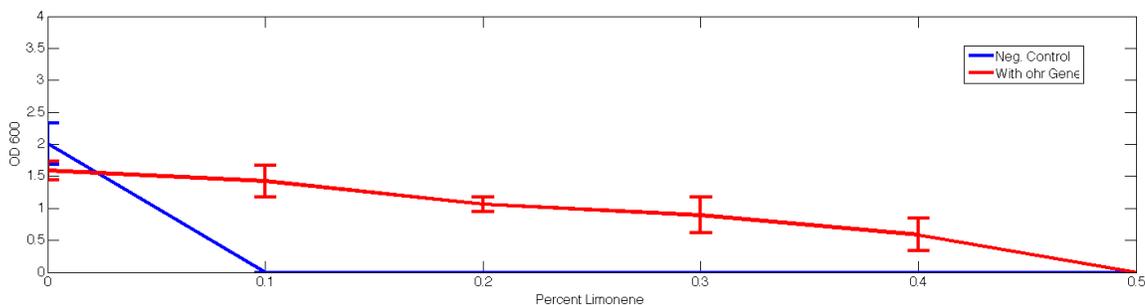


Figure 7: Testing the tolerance range of pBbA5k-ohr in *E. coli* MG1655 versus the negative control (pBbA5k-rfp). Both samples were induced with 100 μ M of IPTG. Limonene was added at time zero and OD600 measurement were taken after 16 hours of growth. The experiment was performed in triplicate with the error bars representing standard deviation.

Our initial work with the *P. aeruginosa* library, acted as a means of verifying the design of the competition experiment and demonstrated that the approach can be used to identify potential biofuel tolerance genes.

CHAPTER 3: FOSMID LIBRARY DESIGN AND CONSTRUCTION

3.1. Designing a Suitable Fosmid Library for Screening

Moving on from the *P. aeruginosa* library, our next step was to construct a new library from a microbe that thrives in hydrocarbon rich environments. We selected *Marinobacter aquaeolei* to meet this criterion. *M. aquaeolei* was discovered living near an oil-producing well in Vietnam, meaning that the bacterium has most likely evolved mechanisms to survive in harsh, hydrocarbon rich environments (Huu, Denner et al. 1999).

Initial attempts at creating a plasmid library using *M. aquaeolei* genomic DNA were hindered by low numbers of unique clones, even when using cells with high transformation efficiencies. We elected to use fosmids in place of plasmids in an attempt to increase the *M. aquaeolei* genome coverage of our library. Fosmid libraries are similar to plasmid libraries except they contain inserts on the order of 40kb and are inserted into cells by using bacteriophages instead of traditional transformation techniques. For a genomic library to have full coverage of the genome, there must be enough unique clones in the library. Using the Clark-Carbon equation below, based on the assumption that recombinant clones follow a Poisson distribution across the genome, library coverage can be calculated with a selected level of certainty that full coverage has been achieved (Clarke and Carbon 1976). P is the probability of full coverage in the library, f is the fraction of the genome contained on one plasmid or fosmid, and N is the number of clones.

$$P = 1 - (1 - f)^N$$

For *M. aquaeolei*, the genome is 4.33Mb long and assuming a 4kb insert size and 95% certainty of full coverage, approximately 3,200 clones would be needed to create a plasmid library. Fosmid libraries have a number of advantages over plasmid libraries. Larger insert sizes (~40kb) drastically reduce the number of clones needed to obtain full coverage by a factor of 10 (> 320 clones). Fosmids also offer good transformation efficiencies (>10⁹ cfu/mL), which can make it easy to get the necessary clones for full coverage. There are however drawbacks to using fosmids in place of plasmids. With such a high number of genes being carried, there is a higher chance that some detrimental heterologous genes may be introduced and effectively negate any potential benefit from positive genes found on the same fosmid. However, work has been done in the past to show that screening fosmids for genes that benefit the host cell can be successful (Ruegg, Kim et al. 2014). With this in mind, we chose fosmids over plasmids to create the *M. aquaeolei* library. Figure 8 illustrates the process of creating the fosmid library.

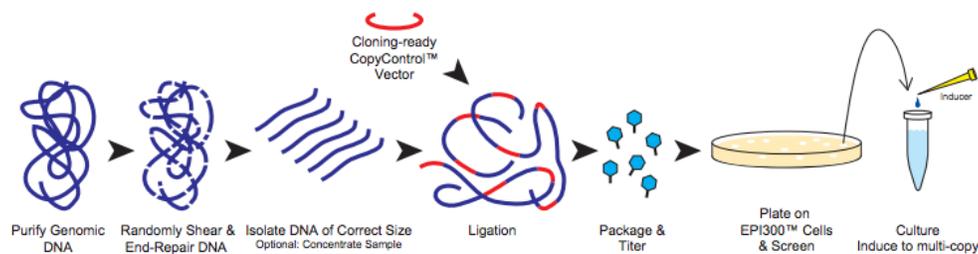


Figure 8: Process of creating the fosmid library (taken from the Epicentre Fosmid Library Production Kit CCFOS110)

To create the library, we used the Epicentre CCFOS110 Fosmid Library Production Kit and were able to generate approximately 1,000 clones. The *M. aquaeolei* genomic DNA used for the creation of the library was obtained from ATCC (#700491). We first ran our genomic DNA on a gel to determine its length. Figure 9 shows that we were starting roughly with genomic DNA fragments on the order of 42 kb. According to the kit protocol, this was an ideal length to use directly as the fosmid insert in the library building process.

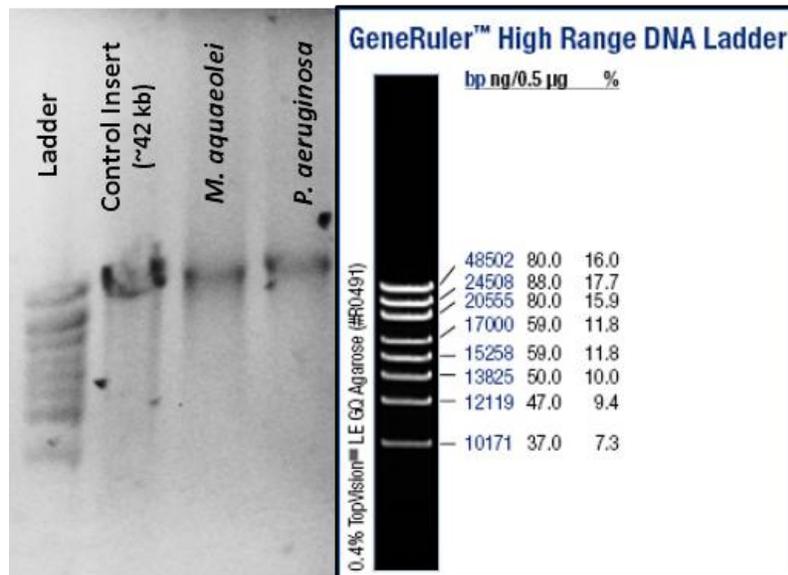


Figure 9. Comparison of genomic DNA samples of versus the human DNA control (~42 kb) provided in the kit and a high range DNA ladder. The genomic DNA for *M. aquaeolei* was centered around ~42 kb meaning that it could be used directly as the insert without gel extraction.

3.2. Testing the *M. aquaeolei* Fosmid Library in Pinene

The *M. aquaeolei* library was then stressed with the bio-jet fuel precursor pinene to determine if any of the *E. coli* harboring the fosmids displayed improved tolerance. The *M. aquaeolei* fosmid library was screened in 0.05% (v/v) pinene over the course of 96 hours, using serial dilutions into fresh media with pinene every 12 hours. We selected this pinene concentration because it severely inhibited growth of the negative control, while cultures with the library survived (Figure 10).

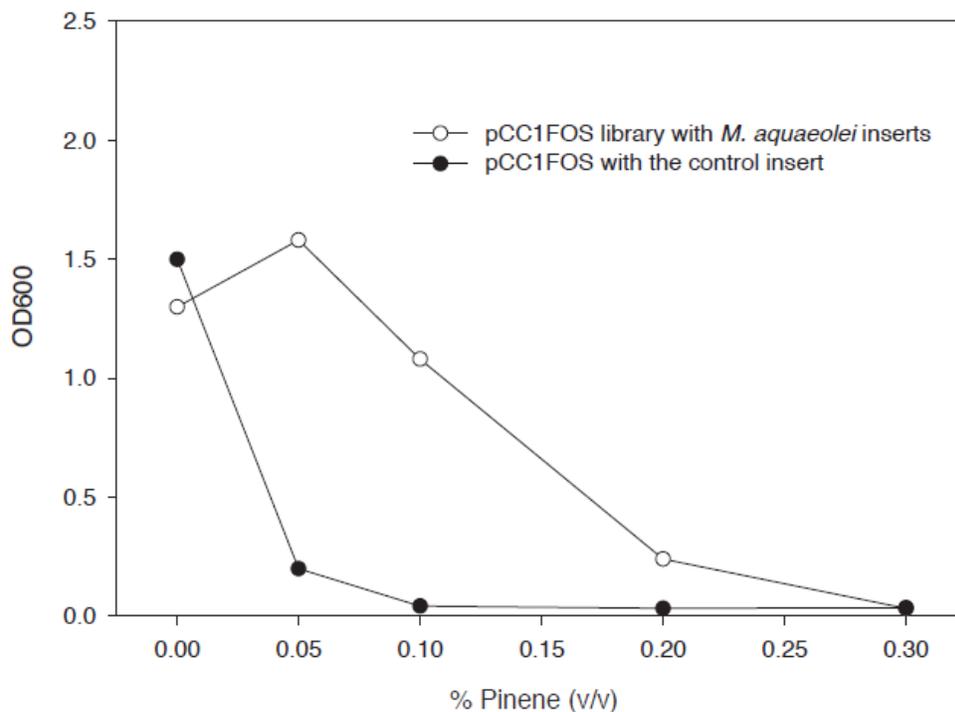


Figure 10. Initial testing of *E. coli* EPI300-TI cells containing the control fosmid and the *M. aquaeolei* fosmid library. 0.05% pinene was selected for subsequent experiments. Measurements of the optical density of the culture were taken at 600nm (OD600).

Based on preliminary tolerance tests, we chose to run the competition experiment (Figure 11) at 0.05% pinene, which inhibited growth of the negative control. When running the competition experiment with the *M. aquaeolei* library we added another control beyond what we used in the *P. aeruginosa* library pilot test. The first control was the fosmid library run with no biofuel added, which controlled for the natural competition between cells to account for any spontaneous mutations that arise. In addition to the fosmid library with no biofuel, we also ran a pBbA5k-rfp negative control with pinene added. The reason for testing the *rfp* control with pinene was to ensure there were no genomic mutations occurring that were causing the improved tolerance. Any observed tolerance increases would have to be coming from the fosmid and not genomic mutations.

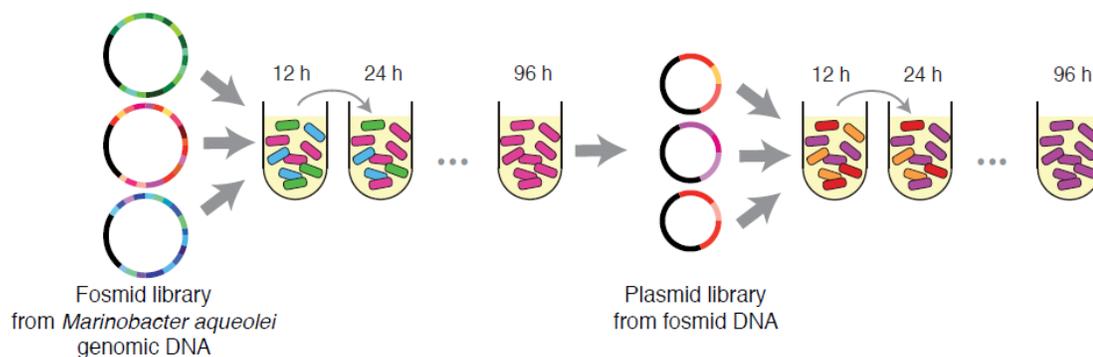


Figure 11. Genomic library approach and sequences isolated in competition experiments. (A) Illustration of the competition experiment used to isolate tolerance genes from *M. aquaeolei*. *E. coli* containing fosmids with *M. aquaeolei* inserts were grown in the presence of 0.05% (v/v) pinene. Every 12 hours, the cultures were diluted into fresh media with pinene. The fosmid that survived the initial competition

experiment was then used as a template for a plasmid library and the competition procedure was repeated.

At the 0, 24, 48, and 96-hour time points we plated cells to isolate single colonies, extracted fosmids, and sequenced to check for library convergence. As expected, at 0 hours all three samples that were sequenced were unique since no stress had been applied to the library yet (Table 1).

Time Point	Sample #	Start	End	Length
0 hours (Fosmid)	1	1,930,787	1,970,656	39.9 kb
	2	2,608,132	2,643,666	35.5 kb
	3	3,177,360	3,219,568	42.2 kb
24 hours (Fosmid)	1	2,535,901	2,567,306	31.4 kb
	2	962,181	1,001,405	39.2 kb
	3	1,881,833	1,915,757	33.9 kb
48 hours (Fosmid)	1	1,885,857	1,926,221	40.4 kb
	2	1,885,854	1,926,222	40.4 kb
	3	1,885,857	1,926,222	40.4 kb
96 hours (Fosmid)	1	1,885,855	1,926,241	40.4 kb
	2	1,885,857	1,926,235	40.4 kb
	3	1,885,856	1,926,237	40.4 kb
96 hours (Plasmid)	1	1,885,857	1,888,294	2.6 kb
	2	1,885,857	1,888,294	2.6 kb
	3	1,885,857	1,888,296	2.6 kb

Table 1. Sequencing results from single colonies at various time points in the selection experiments. Fosmids and plasmids were sequenced using primers listed in Methods and the start and stop position on the *Marinobacter aquaeolei* genome were determined from these results. The numbers in the table correspond to the base pair locations in the GenBank sequence for the *Marinobacter aquaeolei* VT8 complete genome (GenBank: CP000514).

Again, after 24 hours, all sequenced samples were unique with no overlap between the inserts. By 48 hours the three sequenced samples were identical and matched those that were subsequently extracted at the 96-hour time point. We note that one of the three fosmids sequenced from the 24-hour time point overlapped the fosmid sequence from 48 and 96 hours, but the start and end points of the sequence differed. We refer to the converged fosmid here as pCC1FOS-96. The insert contained in pCC1FOS-96 had 43 complete genes from *M. aquaeolei*. The next step was to determine which gene or genes were conferring pinene tolerance to *E. coli*.

3.3. Comparing the competition and the modeling results

From the modeling results presented in section 2.2 it was expected that 96 hours would be required for the library to converge to cells that contain a single shared fosmid. We saw in Table 1 that after 24 hours the three fosmids selected from the competition experiment contained different fosmids. From the model, we can determine

the probability of selecting the best performing fosmid when sampling from the culture after 24 hours $P(\text{pCC1FOS-96}) = 0.029$. In this case, the model matches well with the experimental data because the probability of selecting the fosmid pCC1FOS-96 at the 24 hour time step was very low and none of the samples sequenced at that time step were pCC1FOS-96.

At the 48 hour time step, all three of the sequenced fosmids from the competition were identical. Looking at the model, the probability of selecting pCC1FOS-96 at the 48 hour time step was $P(\text{pCC1FOS-96}) = 0.068$, also a relatively low probability. In fact, the probability that all three randomly selected fosmids were pCC1FOS-96 is $P = 0.068^3 = 0.00031$. Based on what the model showed, we were not expecting to see convergence to a single fosmid at this point.

The model showed convergence happening at a later time step. This is because the standard deviation of the d_i values (the different growth rates of the strains in the library) was too low. What happened in the experiment was that after 48 hours there was enough of a competitive advantage for the cells containing the pCC1FOS-96 fosmid for them to begin dominating the culture. Having a higher standard deviation in the model would account for the fact that some cells in the library have an extreme competitive advantage over the others. The change made to better match the model to our experimental results is illustrated in Figure 12.

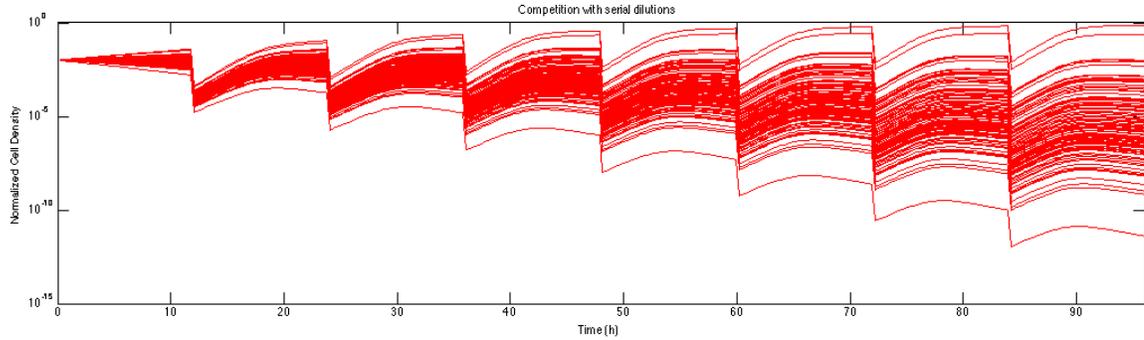


Figure 12. Showing an updated version of the model where the mean of the d_i distribution = 0.75 and the standard deviation has been changed to 0.05.

In this case, the model has clear “winners” of the competition at a much earlier time step (~48 hours), similar to what we observed experimentally.

3.4. Creating and Testing a Sub-library from pCC1FOS-96

To efficiently screen for beneficial genes, we created a sub-library using the fosmid pCC1FOS-96 as the starting template. pCC1FOS-96 DNA was partially digested using the restriction enzyme Sau3A1. Cut sites for Sau3A1 (GATC) appear frequently in the sequence, so we used a partial digest, controlling the length of the reaction and the units of enzyme added. We extracted DNA in the 4-12kb range, selecting this size to allow for complete genes or the possibility of multiple genes without creating inserts that would be prohibitive for cloning or those that return so many genes that it would be difficult to determine contributions. We cloned the inserts

into the medium copy vector pBbA5k (Lee, Krupa et al. 2011), using *E. coli* MG1655 for all subsequent experiments. In order to screen for genes that conferred tolerance, we repeated the competition procedure, stressing cells with 0.05% pinene over the course of 96 hours, as we did with the fosmid library (Figure 12). We conducted the same control experiments as in the fosmid library, replacing the negative control fosmid with pBbA5k-rfp, which expresses red fluorescent protein and does not confer pinene tolerance. Upon completion of the competition experiment, all colonies sampled contained identical plasmids, which we refer to as pBbA5k-96. We extracted and retransformed this plasmid into fresh *E. coli* MG1655 cells to control for the possibility that improvements in tolerance were due to genomic mutations, however the freshly transformed cells retained improved tolerance. This indicated that the presence of the pBbA5k-96 plasmid was responsible for conferring tolerance.

3.5. YceI Improves Pinene Tolerance in *E. coli*

The plasmid pBbA5k-96 contains two complete genes from *M. aquaeolei* with additional truncated genes on each side (Figure 13). The complete genes encode for

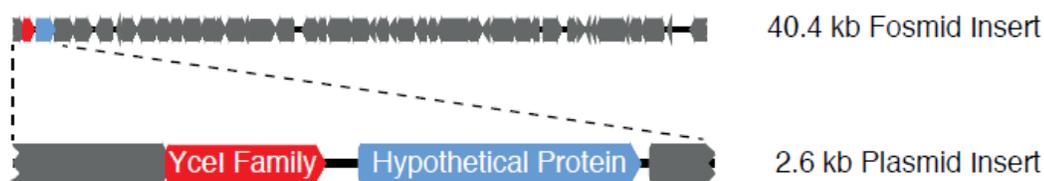


Figure 13: The isolated fosmid and plasmid inserts from the *M. aqueolei* competition experiment

a YceI-family protein, Maqu_1680 [GenBank:YP_958951.1], and a hypothetical protein, Maqu_1681 [GenBank:YP_958952.1]. To determine which gene or genes were responsible for providing increased pinene tolerance, we subcloned the two complete genes individually and together into pBbA8k medium copy vectors. The plasmid contains an arabinose-inducible P_{BAD} promoter, allowing us to control the level of gene expression. Using 100 μ M arabinose we tested the tolerance of cells containing plasmids with the individual and combined genes (Figure 14). Under no pinene stress, cells with YceI grew comparably to the negative control, while all cells expressing the hypothetical protein (alone, in combination with YceI, and on the original pBbA5k-96) experienced a reduction in growth. When exposed to 0.15% pinene, only cells with the YceI protein showed improvements in growth relative to the negative control. Statistical analysis was performed in GraphPad Prism using a one-way ANOVA with a Dunnett test to compare each group of samples with the control in both cases (with and without biofuel). Our tests show that there is an extremely significant benefit to having YceI in the samples when pinene is present (p-value < 0.001). The Dunnett test also showed that there was no statistical benefit for having the Hypothetical Protein.

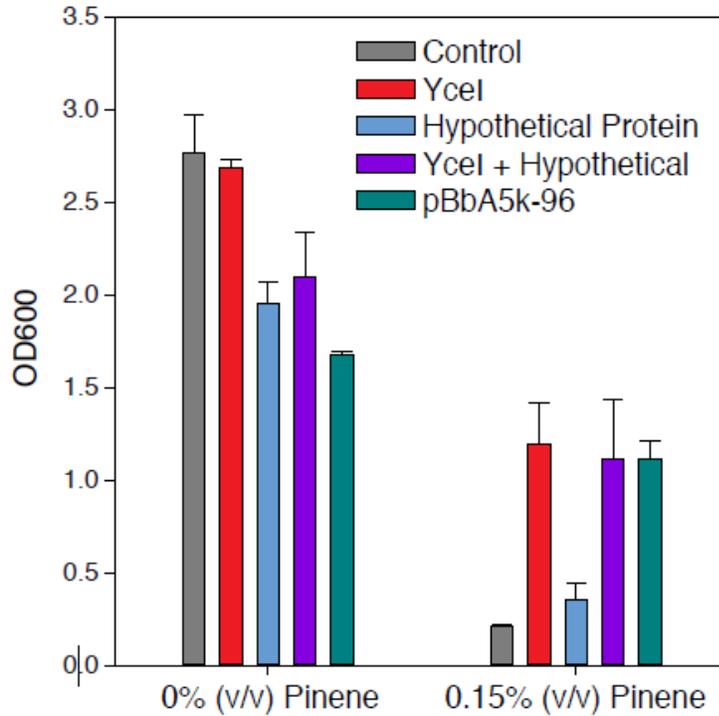


Figure 14. Testing the tolerance of cells containing plasmids with the individual genes identified in the pinene tolerance selection. Measurements of the cell densities were taken at 600nm (OD600). The control is pBbA8k-rfp; other plasmids use the pBbA8k vector and contain either *yceI*, the gene encoding the hypothetical protein, or both. pBbA5k-96 contains both genes as well as the truncated genes at the start and end of the plasmid insert. Error bars represent one standard deviation from the mean.

All cells expressing YceI—including those with YceI alone, in combination with the hypothetical protein, and on pBbA5k-96—showed improvements in tolerance and were statistically significant (p-value < 0.001, one-way ANOVA with Dunnett test). All cells expressing YceI—including those with YceI alone, in combination with the hypothetical protein, and on pBbA5k-96—showed improvements in tolerance and were statistically significant (p-value < 0.001, one-way ANOVA with Dunnett test).

Therefore, we concluded that *yceI* was the sole gene responsible for the increased survival during the competition experiment as well as the increased tolerance observed at 0.15% pinene versus the control.

YceI-family proteins are diverse and remain largely uncharacterized. The YceI-like protein TT1927B from *Thermus thermophiles* HB8 functions as an isoprenoid transport or storage protein and may serve as a part of an unknown isoprenoid metabolic pathway (Handa, Terada et al. 2005). The crystal structure for the protein was solved with the protein complexed with its ligand, a C₄₀ isoprenoid. HP1286 from *Helicobacter pylori* is another YceI-like protein and is overexpressed in acid stress (Sisinni, Cendron et al. 2010). It binds to amphiphilic compounds containing approximately 22 carbon atoms. HP1286 is secreted by the cell, potentially to sequester and supply fatty acids from the environment for metabolism or to detoxify and protect the cell from the antimicrobial properties of fatty acids. *E. coli* harbors a periplasmic YceI-like protein that responds to basic, high pH conditions, but its function has not otherwise been characterized (Stancik, Stancik et al. 2002). YceI from *M. aquaeolei* shares 36% identity with TT1927b, a 31% identity with HP1286, and a 35% identity with the YceI-like protein from *E. coli*. Previously studied proteins in the YceI family have a multitude of functions, ranging from a potential role in the isoprenoid metabolic pathway to involvement in acid or base-induced stress.

We next tested different concentrations of pinene to determine the highest level cells expressing *yceI* could withstand (Figure 15). For all tested pinene levels up to 0.25%, cells expressing *yceI* grew better than the control.

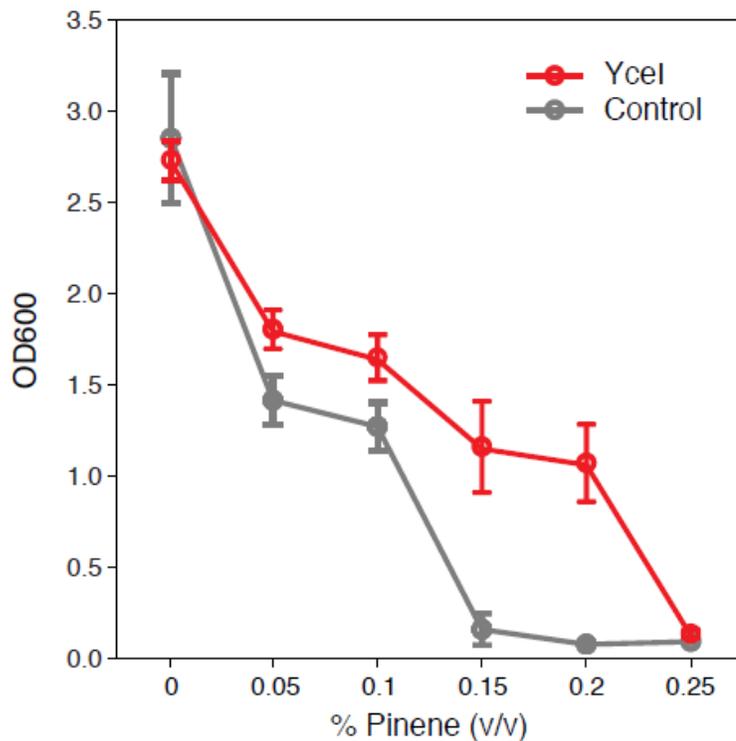


Figure 15. Final OD values under conditions with increasing pinene levels with measurements being taken after 12 hours of growth. Error bars represent one standard deviation from the mean.

We also tested whether overexpression of *yceI* was toxic to cells. Other biofuel tolerance mechanisms, such as efflux pumps are known to provide a benefit under stress, but are toxic if overexpressed, leading to a trade off in expression levels for optimal cell survival (Turner and Dunlop 2014) . We tested varying levels of inducer ranging from 0 to 1000 μ M arabinose and observed no growth impact compared to the control at any concentration of inducer (Figure 16).

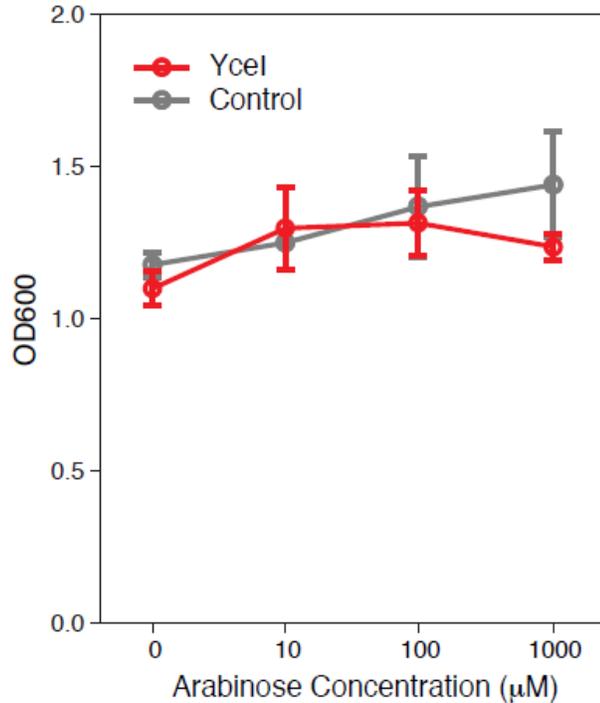


Figure 16. Measuring the toxicity of *yceI* expression compared to the control by inducing expression with arabinose. Error bars represent one standard deviation from the mean.

For all concentrations of arabinose, growth of cells expressing the *yceI* gene were shown through another one-way ANOVA test to be statistically equivalent to the corresponding negative control at each arabinose level.

Due to the similar chemical structure of pinene to the bio-jet fuel precursors terpinolene, terpinene, and limonene (Brennan, Turner et al. 2012), we decided to test tolerance using these chemicals as well. Cells expressing *yceI* were stressed in varying concentrations of the three monoterpene fuels. Our results indicate that *yceI* was beneficial in improving growth when cells were exposed to terpinolene (Figure 17A).

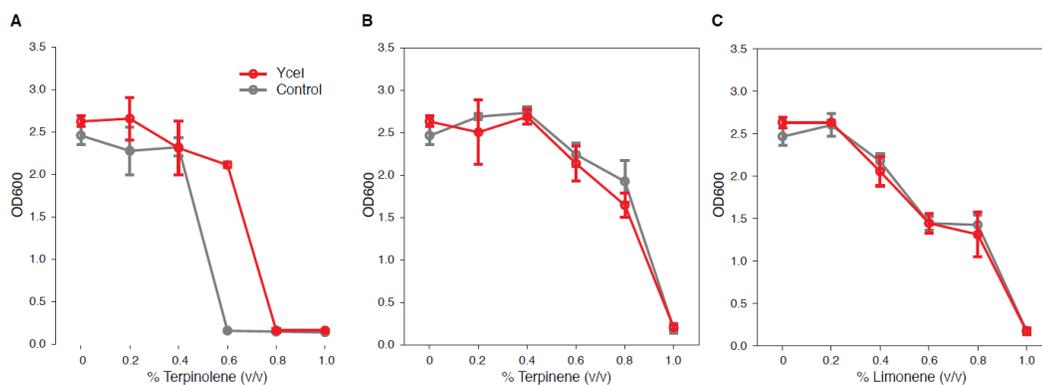


Figure 17. Testing the tolerance effects of *yceI* in the bio-jet fuel precursors (A) terpinolene, (B) terpinene, and (C) limonene. Error bars correspond to one standard deviation from the mean, measured 12 hours after the addition of biofuel.

However, for the two other fuel precursors tested, limonene and terpinene, we saw no significant improvement over the control (Figures 17B-C). It was surprising that we saw such differences in the tolerance benefit of YceI between the four fuels considering that all four of the fuels are monoterpenes ($C_{10}H_{16}$). We began speculating why that might be the case. First of all, looking at the chemical structures of the four fuels tested (Figure 18), the arrangement and locations of the double bonds are different which could change the solubility of these chemicals in water. The solubility in water of isoprenoids, specifically limonene and pinene, has been investigated in the past (Fichan, Larroche et al. 1999). They found that the solubility of limonene in water, although low, is approximately four times greater than that of pinene.

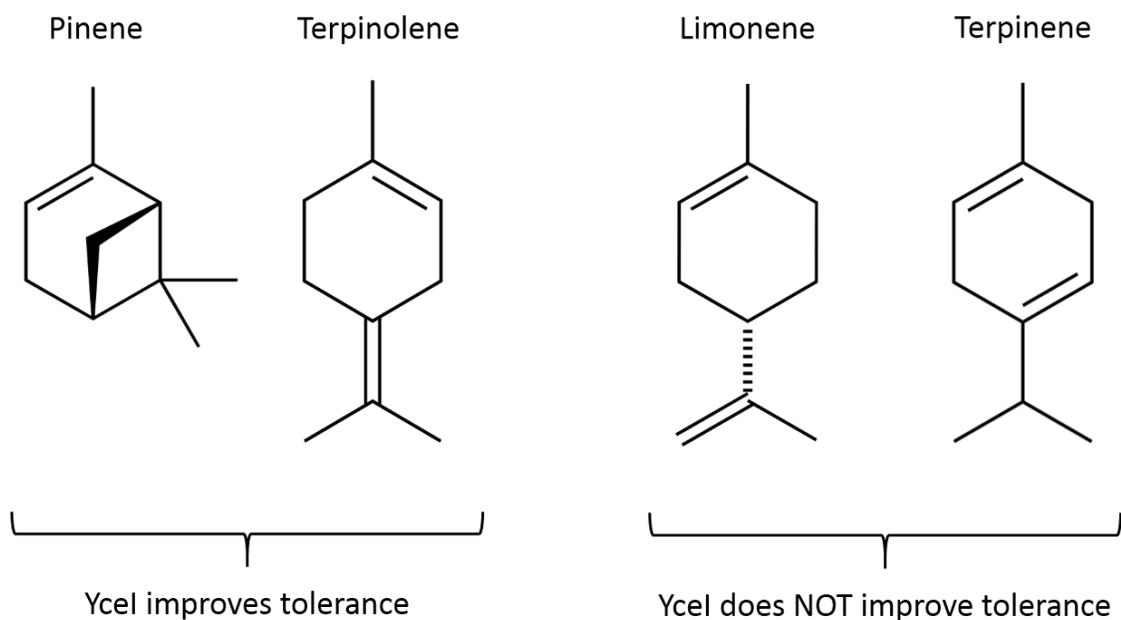


Figure 18. Chemical structures of the four monoterpenes tested in our experiments.

Based on the increased solubility of limonene in water compared to pinene, it is possible that this difference in solubility could be linked to the effectiveness of Ycel in mitigating the stress imparted by the fuel.

CHAPTER 4: FUTURE DIRECTIONS

4.1. Further Work to be Considered

In this study, we used a library approach to screen the genomes of two microbes that have adapted to survive in harsh environments, *P. aeruginosa* and *M. aquaeolei*, for genes that impart tolerance to potential bio-jet fuel precursors. We showed that expression of the gene *ohr* from *P. aeruginosa* and *yceI* from *M. aquaeolei* successfully increased the tolerance of *E. coli* MG1655 to limonene and pinene/terpinolene respectively.

In the future it would be interesting to test both *ohr* and *yceI* in limonene and pinene production strains, respectively, to determine the potential titer increase that could be achieved. Also, combining these two biofuel tolerance genes with other known tolerance mechanisms has the potential to synergistically increase cell survival at higher biofuel concentrations. This has been shown in the past to be a successful strategy for boosting tolerance to higher levels than could be achieved by one mechanism alone (Turner and Dunlop 2014).

This method of library construction and competition could also be used to discover additional tolerance genes. Other hydrocarbon resistant microbes such as *A. borkumensis* or the strains isolated during the Deepwater Horizon disaster, would make excellent sources of genetic material for library construction. Alternatively, instead of selecting individual microbes for library screening, a metagenomic approach could be

applied. Gathering samples of microbial populations from the environment has the potential to provide a diverse sample of genomic DNA.

Given the importance of liquid fuels in our society, it is critical to ensure that we have reliable and economical sources of these fuels for the future. Identifying better ways to produce advanced biofuels will help to lower the cost of their production and make biofuels more competitive with petroleum-derived fuels in the future.

CHAPTER 5: METHODS

5.1. Fosmid Library Construction

Marinobacter aquaeolei VT8 genomic DNA was obtained from ATCC (#700491D-5) and was used for the construction of the fosmid library. The degree of genomic DNA fragmentation was determined via gel electrophoresis on a 0.7% agarose gel run for 20 hours at 30 volts. The gel was then stained and imaged. The genomic DNA had an average length of ~40 kb, which we used directly as the insert in the library construction process. The fosmid library was created using the CopyControl Fosmid Library Production Kit with the pCC1FOS vector (Epicentre). A control fosmid was also created using the ~42 kb fragment of human X-chromosome DNA provided as a control in the kit. *E. coli* EPI300-T1 cells were mixed with the prepared phage particles at different concentrations to determine an appropriate titer, and we selected a 1:10 dilution of the phage particles. The *M. aquaeolei* library and control fosmid were plated on Luria Bertani (LB) agar plates containing 12.5 µg/ml of chloramphenicol, yielding approximately 1,000 colonies per plate after overnight growth at 37C. The colonies containing fasmids with *M. aquaeolei* DNA were scraped off the plate using a razor blade and 50 µl aliquots of the library were stored in a 20% glycerol solution at -80C for later use. An individual colony from the control fosmid plate was used to create a glycerol stock.

5.2. Selection Procedure: Fosmid Library

LB supplemented with 12.5 µg/ml of chloramphenicol was used to prepare two cultures containing *E. coli* EPI300-T1 cells harboring the *M. aquaeolei* fosmid library. A third culture containing the human DNA control fosmid in *E. coli* EPI300-T1 was prepared in a similar manner. One of the *M. aquaeolei* library cultures and the control were stressed with 0.05% pinene (v/v), while the other *M. aquaeolei* library culture received no exposure to pinene. Each of the cultures was diluted 1:100 into LB plus chloramphenicol and pinene, where applicable, every 12 hours over the course of a 96 hour period. Every 24 hours, samples from each culture were plated on LB agar plates with chloramphenicol. Three colonies from each of the 0, 24, 48, and 96-hour time points were selected and their fosmids were extracted as follows: Cells were grown in 5ml of LB medium plus chloramphenicol and 10µl of 500X CopyControl Fosmid Autoinduction solution from the Epicentre Kit. The culture was grown for 16 hours and the fosmid DNA was then extracted using a Qiagen QIAprep Spin Miniprep Kit. The fosmid samples sequenced using the pCC1 forward (5'-GGATGTGCTGCAAGGCGATTAAGTTGG - 3') and reverse (5'-CTCGTATGTTGTGTGGAATTGTGAGC - 3') primers indicated in the Epicentre Kit protocol. The resulting sequences were aligned with the *M. aquaeolei* genome. The converged fosmid from the 96-hour time point, which we refer to as pCC1FOS-96, was saved for later use.

5.3. Plasmid Library Construction

The fosmid from the converged 96-hour time point (pCC1FOS-96) was used as the starting material for plasmid library construction. 500 µg of the fosmid DNA was partially digested at 37C using varying concentrations (0.13U, 0.25U, 0.38U, and 0.50U) of the enzyme Sau3A1 (New England Biolabs). By reducing the digestion time to five minutes, we achieved a DNA insert length centered around 4-12 kb. The digested fosmid DNA from each of the four reactions was then gel extracted from the 4-12 kb range using the Qiagen QIAquick Gel Extraction Kit (Wargo, Szwergold et al. 2008). We used pBbA5k (Lee, Krupa et al. 2011) as the plasmid vector, which has a medium copy p15A origin of replication, lacUV5 promoter upstream of the cloning site, and a kanamycin resistance gene. To prepare the vector, the pBbA5k-rfp plasmid was double digested using BamHI and BglII and gel extracted. The pBbA5k vector and prepared library inserts were ligated using T4 DNA ligase (Fermentas) at a 3:1 insert to vector ratio. The ligated mixture was transformed into *E. coli* MG1655 and plated onto LB agar plates containing 50 µg/ml kanamycin. Plates were incubated overnight at 37C and the resulting ~500 colonies were scraped off of the plates using a razor blade and stored at -80C in a 20% glycerol solution, as described above.

5.4. Selection Procedure: Plasmid Library

The selection experiment run using the plasmid library was the same as the fosmid selection described above, with the following changes: pBbA5k-rfp in *E. coli* MG1655 was used as the control, 50 µg/ml of kanamycin was used in place of chloramphenicol, and sequencing was performed on 3 samples from the pinene-treated plasmid library plate at 96 hours using the forward (5'-GGAATTGTGAGCGGATAACAATTC-3') and reverse (5'-CGTTTTATTTGATGCCTGGAGATCC-3') primers for the pBbA5k vector, as given in (Lee, Krupa et al. 2011). The converged to plasmid after the 96 hours was saved for later use and named pBbA5k-96.

5.5. Subcloning Genes from the Converged Plasmid

Two *M. aquaeolei* genes encoding for the YceI family protein [GenBank:YP_958951.1] and the hypothetical protein [GenBank:YP_958952.1] were subcloned using the pBbA8k BioBrick vector (Lee, Krupa et al. 2011), which has a p15A origin of replication, P_{BAD} promoter, and kanamycin resistance cassette. The DNA fragments were cloned using the Gibson Assembly Protocol (Gibson, Young et al. 2009). Inserts were prepared using PCR and the vector was prepared by digesting pBbA8k-rfp with BamHI and BglIII and gel extracting. Individual colonies of each construct were isolated and cloning success was verified via sequencing using the

forward (5'-CTACTGTTTCTCCATACCCGTTTTTTTGG-3') and reverse (5'-CGTTTTATTTGATGCCTGGAGATCC-3') primers for the pBbA8k vector (Lee, Krupa et al. 2011).

5.6. Tolerance Testing Procedure

Overnight cultures were grown for 16 hours in LB containing antibiotics and varying levels of arabinose (as required). Cultures for tolerance testing were prepared by first preparing a pre-culture, where we inoculated 5 ml of LB containing antibiotics and arabinose (as required) with 50 μ L of the overnight culture. Cultures were grown until they reached an OD600 reading of 0.2, at which point varying levels of biofuel were added to each. The chemicals used for tolerance testing were obtained from Sigma Aldrich (α -pinene P45680, γ -terpinene 86478, limonene 183164, and terpinolene W304603). Growth measurements were taken after 12 hours under biofuel stress. All experiments were performed in triplicate.

References:

Alonso-Gutierrez, J., R. Chan, T. S. Bath, P. D. Adams, J. D. Keasling, C. J. Petzold and T. S. Lee (2013). "Metabolic engineering of *Escherichia coli* for limonene and perillyl alcohol production." Metabolic Engineering **19**: 33-41.

Alper, H., J. Moxley, E. Nevoigt, G. R. Fink and G. Stephanopoulos (2006). "Engineering yeast transcription machinery for improved ethanol tolerance and production." Science **314**(5805): 1565-1568.

Alsaker, K. V., C. Paredes and E. T. Papoutsakis (2010). "Metabolite stress and tolerance in the production of biofuels and chemicals: Gene-expression-based systems analysis of butanol, butyrate, and acetate stresses in the anaerobe *Clostridium acetobutylicum*." Biotechnol Bioeng: n/a-n/a.

Atsumi, S., T.-Y. Wu, I. M. P. Machado, W.-C. Huang, P.-Y. Chen, M. Pellegrini and J. C. Liao (2010). "Evolution, genomic analysis, and reconstruction of isobutanol tolerance in *Escherichia coli*." Mol. Syst. Biol. **6**: 449.

Brennan, T. C. R., C. D. Turner, J. O. Krömer and L. K. Nielsen (2012). "Alleviating monoterpene toxicity using a two-phase extractive fermentation for the bioproduction of jet fuel mixtures in *Saccharomyces cerevisiae*." Biotechnol Bioeng **109**(10): 2513-2522.

Camilli, R., C. M. Reddy, D. R. Yoerger, B. A. S. Van Mooy, M. V. Jakuba, J. C. Kinsey, C. P. McIntyre, S. P. Sylva and J. V. Maloney (2010). "Tracking hydrocarbon plume transport and biodegradation at Deepwater Horizon." Science **330**(6001): 201-204.

Chuck, C. J. and J. Donnelly (2014). "The compatibility of potential bioderived fuels with Jet A-1 aviation kerosene." Appl Energy **118**: 83-91.

Clarke, L. and J. Carbon (1976). "A colony bank containing synthetic Col El hybrid plasmids representative of the entire *E. coli* genome." Cell **9**(1): 91-99.

Doshi, R., T. Nguyen and G. Chang (2013). "Transporter-mediated biofuel secretion." Proceedings Of The National Academy Of Sciences Of The United States Of America **110**(19): 7642-7647.

Dunlop, M. J. (2011). "Engineering microbes for tolerance to next-generation biofuels." Biotechnol Biofuels **4**: 32.

Dunlop, M. J., Z. Y. Dossani, H. L. Szmids, H. C. Chu, T. S. Lee, J. D. Keasling, M. Z.

Hadi and A. Mukhopadhyay (2011). "Engineering microbial biofuel tolerance and export using efflux pumps." Mol. Syst. Biol. **7**: 487.

Fichan, I., C. Larroche and J. Gros (1999). "Water solubility, vapor pressure, and activity coefficients of terpenes and terpenoids." Journal of Chemical & Engineering Data **44**(1): 56-62.

Fiocco, D., V. Capozzi, P. Goffin, P. Hols and G. Spano (2007). "Improved adaptation to heat, cold, and solvent tolerance in *Lactobacillus plantarum*." Appl. Microbiol. Biotechnol. **77**(4): 909-915.

Fischer, C. R., D. Klein-Marcuschamer and G. Stephanopoulos (2008). "Selection and optimization of microbial hosts for biofuels production." Metabolic Engineering **10**(6): 295-304.

Foo, J. L., H. M. Jensen, R. H. Dahl, K. George, J. D. Keasling, T. S. Lee, S. Leong and A. Mukhopadhyay (2014). "Improving microbial biogasoline production in *Escherichia coli* using tolerance engineering." MBio **5**(6): e01932.

Foo, J. L. and S. S. J. Leong (2013). "Directed evolution of an *E. coli* inner membrane transporter for improved efflux of biofuel molecules." Biotechnology For Biofuels **6**(1): 81.

Fortman, J. L., S. Chhabra, A. Mukhopadhyay, H. Chou, T. S. Lee, E. Steen and J. D. Keasling (2008). "Biofuel alternatives to ethanol: pumping the microbial well." Trends Biotechnol. **26**(7): 375-381.

Gibson, D. G., L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison and H. O. Smith (2009). "Enzymatic assembly of DNA molecules up to several hundred kilobases." Nat Methods **6**(5): 343-345.

Handa, N., T. Terada, Y. Doi-Katayama, H. Hirota, J. R. H. Tame, S.-Y. Park, S. Kuramitsu, M. Shirouzu and S. Yokoyama (2005). "Crystal structure of a novel polyisoprenoid-binding protein from *Thermus thermophilus* HB8." Protein Sci. **14**(4): 1004-1010.

Harvey, B. G., M. E. Wright and R. L. Quintana (2010). "High-Density Renewable Fuels Based on the Selective Dimerization of Pinenes." Energy Fuels **24**(1): 267-273.

Hazen, T. C., E. A. Dubinsky, T. Z. DeSantis, G. L. Andersen, Y. M. Piceno, N. Singh, J. K. Jansson, A. Probst, S. E. Borglin, J. L. Fortney, W. T. Stringfellow, M. Bill, M. E. Conrad, L. M. Tom, K. L. Chavarria, T. R. Alusi, R. Lamendella, D. C. Joyner, C. Spier, J. Baelum, M. Auer, M. L. Zemla, R. Chakraborty, E. L. Sonnenthal, P. D'haeseleer, H.-Y. N. Holman, S. Osman, Z. Lu, J. D. Van Nostrand, Y. Deng, J.

- Zhou and O. U. Mason (2010). "Deep-Sea Oil Plume Enriches Indigenous Oil-Degrading Bacteria." Science (New York, NY) **330**(6001): 204-208.
- Hong, M.-E., K.-S. Lee, B. J. Yu, Y.-J. Sung, S. M. Park, H. M. Koo, D.-H. Kweon, J. C. Park and Y.-S. Jin (2010). "Identification of gene targets eliciting improved alcohol tolerance in *Saccharomyces cerevisiae* through inverse metabolic engineering." Journal of Biotechnology **149**(1-2): 52-59.
- Huu, N. B., E. B. Denner, D. T. Ha, G. Wanner and H. Stan-Lotter (1999). "Marinobacter aquaeolei sp. nov., a halophilic bacterium isolated from a Vietnamese oil-producing well." Int J Syst Bacteriol **49 Pt 2**: 367-375.
- Lee, S. K., H. Chou, T. S. Ham, T. S. Lee and J. D. Keasling (2008). "Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels." Current Opinion in Biotechnology **19**(6): 556-563.
- Lee, T. S., R. A. Krupa, F. Zhang, M. Hajimorad, W. J. Holtz, N. Prasad, S. K. Lee and J. D. Keasling (2011). "BglBrick vectors and datasheets: A synthetic biology platform for gene expression." Journal of Biological Engineering **5**(1): 12.
- Lee, T. S., R. A. Krupa, F. Zhang, M. Hajimorad, W. J. Holtz, N. Prasad, S. K. Lee and J. D. Keasling (2011). "BglBrick vectors and datasheets; a synthetic biology platform for gene expression." Journal of Biological Engineering **5**(1): 12.
- Lesniak, J., W. A. Barton and D. B. Nikolov (2002). Structural and functional characterization of the *Pseudomonas* hydroperoxide resistance protein Ohr. EMBO J, **21**: 6649-6659.
- Li, X. Z., L. Zhang and K. Poole (1998). "Role of the multidrug efflux systems of *Pseudomonas aeruginosa* in organic solvent tolerance." J. Bacteriol. **180**(11): 2987-2991.
- Nicolaou, S. A., S. M. Gaida and E. T. Papoutsakis (2010). "A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation." Metabolic Engineering **12**(4): 307-331.
- Nicolaou, S. A., S. M. Gaida and E. T. Papoutsakis (2012). "Exploring the combinatorial genomic space in *Escherichia coli* for ethanol tolerance." Biotechnol J **7**(11): 1337-1345.
- Peabody, G. L., J. Winkler and K. C. Kao (2014). "Tools for developing tolerance to toxic chemicals in microbial systems and perspectives on moving the field forward and into the industrial setting." Current Opinion in Chemical Engineering.

Ramos, J. L., E. Duque, M.-T. Gallegos, P. Godoy, M. I. Ramos-Gonzalez, A. Rojas, W. Teran and A. Segura (2002). "Mechanisms of solvent tolerance in gram-negative bacteria." Annu. Rev. Microbiol. **56**: 743-768.

Reyes, L. H., M. P. Almario and K. C. Kao (2011). "Genomic library screens for genes involved in n-butanol tolerance in Escherichia coli." PLoS ONE.

Rojas, A., E. Duque, G. Mosqueda, G. Golden, A. Hurtado, J. L. Ramos and A. Segura (2001). "Three efflux pumps are required to provide efficient tolerance to toluene in Pseudomonas putida DOT-T1E." J. Bacteriol. **183**(13): 3967-3973.

Ruegg, T. L., E.-M. Kim, B. A. Simmons, J. D. Keasling, S. W. Singer, T. S. Lee and M. P. Thelen (2014). "An auto-inducible mechanism for ionic liquid resistance in microbial biofuel production." Nat Commun **5**: 3490.

Sarria, S., B. Wong, H. García Martín, J. D. Keasling and P. Peralta-Yahya (2014). "Microbial synthesis of pinene." ACS Synth Biol **3**(7): 466-475.

Schneiker, S., V. A. P. M. dos Santos, D. Bartels, T. Bekel, M. Brecht, J. Buhrmester, T. N. Chernikova, R. Denaro, M. Ferrer, C. Gertler, A. Goesmann, O. V. Golyshina, F. Kaminski, A. N. Khachane, S. Lang, B. Linke, A. C. McHardy, F. Meyer, T. Nechitaylo, A. Pühler, D. Regenhardt, O. Rupp, J. S. Sabirova, W. Selbitschka, M. M. Yakimov, K. N. Timmis, F.-J. Vorhölter, S. Weidner, O. Kaiser and P. N. Golyshin (2006). "Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium Alcanivorax borkumensis." Nat Biotechnol **24**(8): 997-1004.

Schobert, H. H. (2013). "The Chemistry of Hydrocarbon Fuels." 356.

Sisinni, L., L. Cendron, G. Favaro and G. Zanotti (2010). "Helicobacter pylori acidic stress response factor HP1286 is a YceI homolog with new binding specificity." FEBS J **277**(8): 1896-1905.

Stancik, L. M., D. M. Stancik and B. Schmidt (2002). "pH-dependent expression of periplasmic proteins and amino acid catabolism in Escherichia coli." Journal of Bacteriology.

Strogatz, S. (2005). Nonlinear dynamics and chaos with applications to physics, biology, chemistry, and engineering. Cambridge, Mass, Westview Press.

Turner, W. J. and M. J. Dunlop (2014). "Trade-Offs in Improving Biofuel Tolerance Using Combinations of Efflux Pumps." ACS Synth Biol.

Wargo, M. J. and D. A. Hogan (2009). "Identification of genes required for Pseudomonas aeruginosa carnitine catabolism." Microbiology **155**(7): 2411-2419.

Wargo, M. J., B. S. Szwegold and D. A. Hogan (2008). "Identification of two gene clusters and a transcriptional regulator required for *Pseudomonas aeruginosa* glycine betaine catabolism." J. Bacteriol. **190**(8): 2690-2699.

Weber, H., T. Polen, J. Heuveling, V. F. Wendisch and R. Hengge (2005). "Genome-Wide Analysis of the General Stress Response Network in *Escherichia coli*: S-Dependent Genes, Promoters, and Sigma Factor Selectivity." J. Bacteriol. **187**(5): 1591-1603.

Woodruff, L. B. A., J. Pandhal, S. Y. Ow, A. Karimpour-Fard, S. J. Weiss, P. C. Wright and R. T. Gill (2013). "Genome-scale identification and characterization of ethanol tolerance genes in *Escherichia coli*." Metabolic Engineering **15**(C): 124-133.

Zingaro, K. A., S. A. Nicolaou, Y. Yuan and E. T. Papoutsakis (2014). "Exploring the heterologous genomic space for building, stepwise, complex, multicomponent tolerance to toxic chemicals." ACS Synth Biol **3**(7): 476-486.