

**REPORT DOCUMENTATION PAGE**Form Approved  
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE (DD-MM-YYYY)</b> 10-09-2013		<b>2. REPORT TYPE</b> Final Technical		<b>3. DATES COVERED (From - To)</b> 6/1/12 - 5/31/13	
<b>4. TITLE AND SUBTITLE</b> Engineering Complex Microbial Phenotypes with continuous genetic integration and plasmid based multi-gene library				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b> N00014-10-1-0161	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b> Papoutsakis, Eleftherios				<b>5d. PROJECT NUMBER</b> 09PR09747-00	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> University of Delaware 210 Hullihen Hall Newark, DE 19716				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b> FINAL	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> Office of Naval Research 875 North Randolph Street Arlington, VA 22203-1995				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b> ONR	
				<b>11. SPONSORING/MONITORING AGENCY REPORT NUMBER</b>	
<b>12. DISTRIBUTION AVAILABILITY STATEMENT</b> Public					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b> In this project we developed imaging systems based on the principles of compressive sensing. We developed advanced signal processing algorithms as well as a hardware testbed. The key idea in CS reconstruction is the realization that most signals encountered in practice are sparse in some domain and the theory of CS exploits such sparsity to dictate that far fewer sampling resources than traditional approaches are needed. Spectral image cubes are particularly well suited for sparse representation as images across different wavelengths exhibit strong correlation. We have thus developed a testbed to acquire and reconstruct natural images from a limited number of linear projection measurements at sub-Nyquist sampling rates. A key to the success of CS is the design of the measurement ensemble, which is based on the evaluation of the incoherence between the measurement ensemble and the sparsity basis. Due to the large scale nature of images, the generation of the measurement ensemble should be both computationally efficient and memory efficient. Based on these principles, we developed compressive cross-sectional imaging (confocal) systems, a single pixel compressive camera, and a multispectral imaging system. Advanced algorithms for the efficient reconstruction of the compressive measurements were developed.					
<b>15. SUBJECT TERMS</b> DNA, ethanol, butanol					
<b>16. SECURITY CLASSIFICATION OF:</b>		<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>	

20140506568

a. REPORT

INSTRUCTIONS FOR COMPLETING SF 298

b. ABSTRACT

19b. TELEPHONE NUMBER (Include area code)

c. THIS PAGE

**PI NAME: Eleftherios Papoutsakis**

**Department of Chemical Engineering and the Delaware Biotechnology Institute,  
University of Delaware**

**ONR award number: N000141010161**

**ONR Award Title: Engineering Complex Microbial Phenotypes with continuous genetic integration and plasmid based multi-gene library**

**FINAL REPORT: October 1, 2009 to March 31, 2013**

**(project had an initial completion date of 9/30/2012, but a no cost extension was requested and granted to extend the project to 3/30/2013)**

## **II. Scientific and Technical Objectives**

Develop tools & strategies to generate complex microbial phenotypes by combining DNA from different organisms aiming to develop desirable complex traits, and notably tolerance to chemicals, such as ethanol (EtOH), butanols (n-BuOH and i-BuOH) and 1,2,4-butanetriol (BT). BT is of interest to Navy for production by a biological process.

**Objective 1:** Build complex phenotypes by accelerated evolution. The goal is to enlarge the *Escherichia coli* genome by selective integration of genes from strains that have desirable traits. We aim to enlarge the *E. coli* genome using *Lactobacillus plantarum* genes to build cells tolerant to EtOH and BT. *L. plantarum* is an organism with established high tolerance to alcohols and solvents.

**Objective 1a:** To further expand our toolbox for genome engineering, we desired to develop library-based screening methods that capture interactions among genetic loci, so that this can be applied to allogeneic DNA (including metagenomic DNA) as well as homologous DNA. We aimed to test the approach first with *E. coli* DNA to screen for ethanol and oxidative stress tolerance.

**Objective 2:** Build a stress-response (or chaperone) system that can be customized for tolerance to chosen toxic chemicals, such as EtOH and BT. Test the improved tolerance strains for improved productivity of industrially relevant solvents.

**Objective 2a (new evolved objective):** Identify new, heterologous genes for stress-response system beyond those from rational design approach using library screenings for improved tolerance to toxic chemicals

## **III. Approach**

**Objective 1 & 1a:** Integrated heterologous (*L. plantarum*) DNA into the *E. coli* chromosome and selected for insertions that enhance ethanol tolerance. In YR1, we used the originally proposed approach, which worked, but not as effectively as anticipated. During YR2, we used a new approach based on the Cre-Lox system which has been used extensively in

eukaryotic systems, but not in bacterial systems. In a synergistic approach, we generated co-existing genomic libraries that can be co-expressed in a single host, and screened the libraries for ethanol and oxidative stress tolerance to determine if beneficial gene interactions can be identified. This was expanded to allogeneic DNA as we discuss below.

**Objective 2 & 2a.** Aiming to develop a semi-synthetic stress response system as a means to build tolerance to toxic chemicals, we aimed to screen the impact of co-expressed HSPs on the development of tolerance. We have also performed ethanol-production assays<sup>1221</sup> using the stress-response system previously developed in combination with genes identified from screening *L. plantarum* libraries. Additionally, we have screened heterologous libraries for genes imparting improved survival and growth in ethanol in combination with the HSP genes we have previously identified.

#### **IV. Concise Accomplishments**

##### **Objective 1 & 1a:**

1. The benefit from a 4288 bp fragment identified from *L. plantarum* genomic integration was determined to be derived from the murA2 gene in combination with a 973 bp upstream non-coding region. Characterization of this region was completed and the murA2 gene was determined to be responsible while the upstream region was identified as a non-coding RNA that appears to interact with an upstream untranslated region in the murA2 promoter.
2. Further characterized identified strains for improved oxidative stress and explored the impact of intracellular iron concentrations on oxidative stress tolerance in combination with identified putative iron transporters.

##### **Objective 2:**

1. Demonstrated improved tolerance in ethanol, n-butanol, and BT when *E. coli* HSPs were overexpressed on single plasmid system
2. Tested combinations of HSP genes and *L. plantarum* murA2 fragment in *E. coli* for improved ethanol production in ethanologenic KO11 strain.
3. Screened heterologous genomic libraries in combination with stress-response system using strains engineered with heterologous transcription machinery for improved tolerance to ethanol for improvements to both survival and growth under ethanol stress.

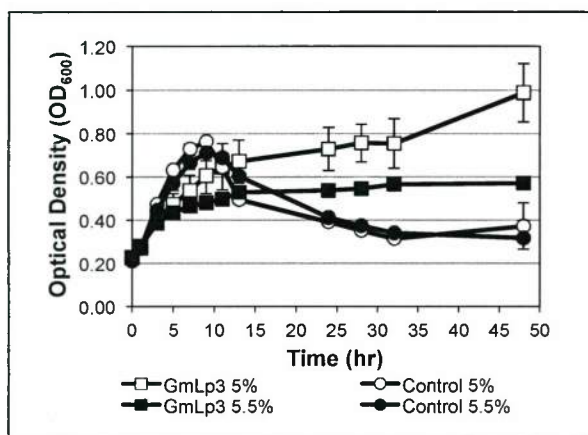
## V. Expanded Accomplishments

### Objective 1

The aim of the project was to demonstrate that heterologous DNA from *L. plantarum* can be incorporated into the *E. coli* chromosome, and then screened for an improved phenotype. The exogenous DNA was incorporated into the chromosome of *E. coli* using modified Cre-Lox sites, as detailed in the previous reports. The inserted *L. plantarum* DNA was flanked with the constitutive *bla* promoter (beta-lactamase) to ensure transcription once incorporated in the *E. coli* chromosome. A collection of *E. coli* strains with *L. plantarum* fragments were generated and screened for ethanol tolerance. A promising strain, designated **EC100GmLp3**, was identified that exhibited improved growth and survival under ethanol stress. The data presented here demonstrate the characterization of the improved phenotype of the strain created, as well as our efforts to fully sequence the inserted DNA.

### Demonstrating the enhanced phenotype of the GmLp3 strain

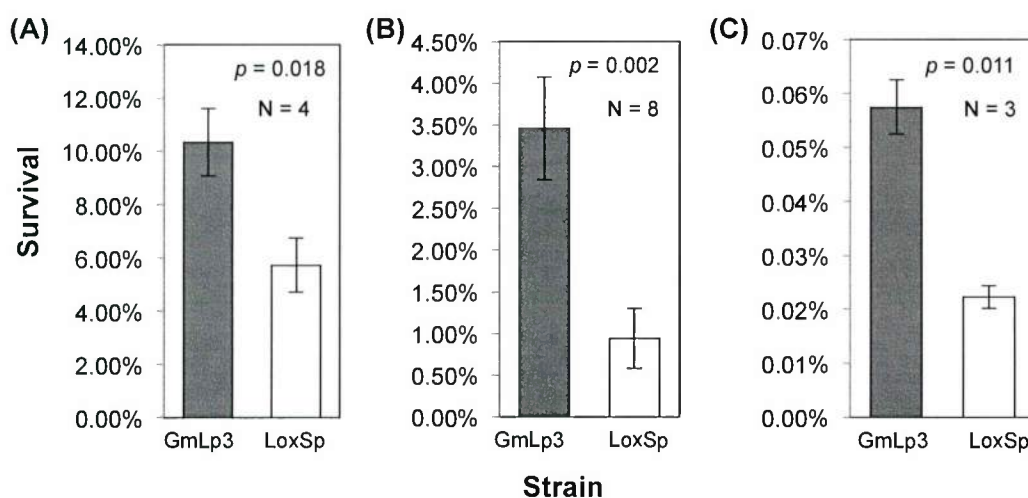
The growth of the EC100GmLp3 and the control LoxSp strains under ethanol stress were investigated to determine if the generated strain had enhanced tolerance in alcohol. It must be noted that the control LoxSp strain is the parent strain with the initial insertion of the spectinomycin resistance gene, whereas the GmLp3 strain is the experimentally selected strain after the library selection, which contains *L. plantarum* fragments integrated in the chromosome. The strains are different, but the parent strain was used as a control to compare the relative changes due to the additional genomic insertions. Fig. 1 demonstrates the growth profiles of the two strains at two ethanol concentrations, 5% and 5.5% (v/v). The GmLp3 strain grows to higher densities under both ethanol concentrations, demonstrating improved tolerance to the stressor.



**Fig. 1:** Growth curves of the control (circles) and the GmLp3 (squares) strains in 5% (white) and 5.5% (black) (v/v) ethanol. The GmLp3 strain, which contains *L. plantarum* DNA incorporated into the *E. coli* chromosome grows to higher optical densities.

In addition to growth in the presence of ethanol, the survival of the two strains under different concentrations of ethanol was tested. This survival assay relied on the determination of the pre- and post-stress colony forming units (CFU) on agar plates. Survival assays give a better indication of the overall survivability of a strain under solvent stress in production-like conditions. Ethanol concentrations of 5%, 6% and 8% (v/v) were utilized. The strains were stressed in the presence of ethanol for a total of 48 hours.

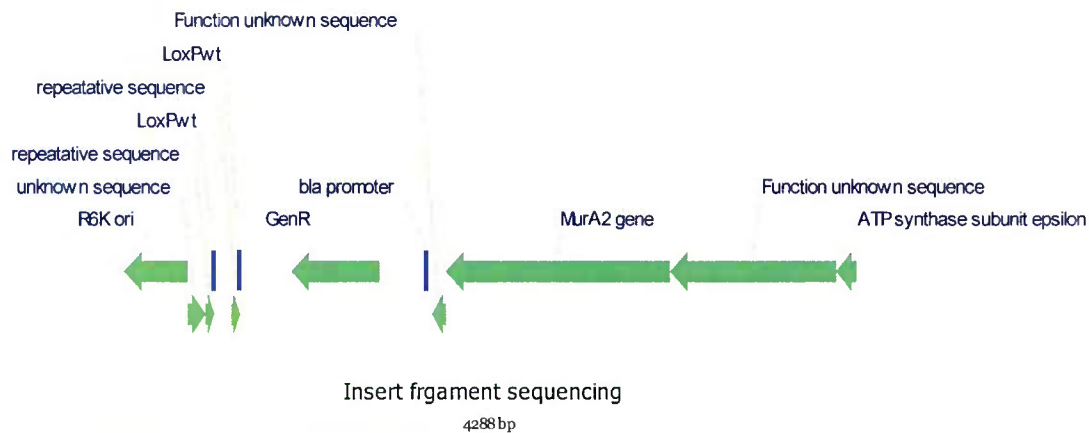
The data from three technical experiments of  $N$  biological replicates, four for 5%, eight for 6% ethanol and three for 8% ethanol, were used to compare the survival rates between the two strains at the different ethanol concentrations, and the results are illustrated in Fig. 2. The GmLp3 strain resulted in more CFUs compared to the control strain LoxSp at all ethanol concentrations tested and had an overall higher survival rate. It is notable that as the ethanol concentration increases from 5% to 6%, the fold improvement of the experimental strain versus the control increases from 1.8-fold to 3.7-fold, which illustrates the higher survival rate of the GmLp3 strain and its ability to survive in 6% ethanol for 48 hours. Furthermore, the GmLp3 strain is better able to survive the stress at 6% ethanol, with only a 3-fold decrease in survival compared to the 5% ethanol stress. The control strain survival rate decreased from 5.7% to 0.94%, or approximately 6-fold, demonstrating that this ethanol concentration is inhibitory to the cells. At the highest ethanol concentration tested (8%) the experimental strain still produced more CFUs, but the overall survival rate was low for both strains. This is expected since due to the high stress levels, but again the experimental strain outperforms the control. In 8% ethanol, the improvement was 2.6-fold.



**Fig. 2:** The percent survival rates of the selected EC100GmLp3 (gray) and the control LoxSp (white) after 48 hours of ethanol stress at different concentrations. (A) 5%, (B) 6% and (C) 8% ethanol (v/v). Error bars illustrate the standard error of  $N$  biological experiments, each with three technical replicates. The  $p$ -value of the t-test comparing the mean survival rate at each ethanol concentration is also shown.

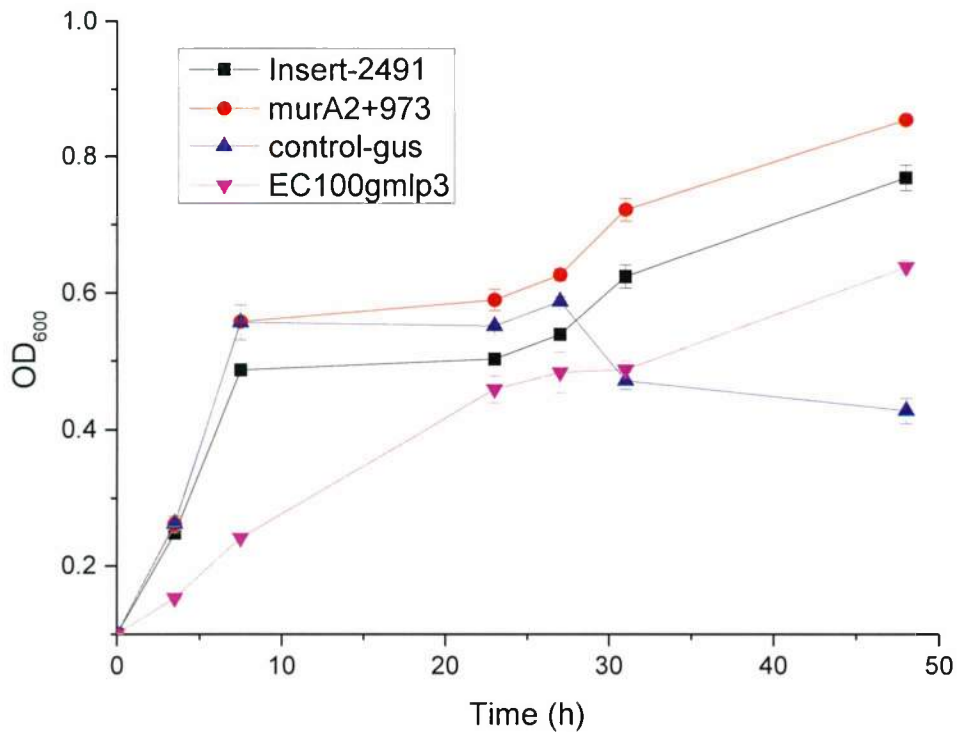
### Sequencing the inserts of the GmLp3 strain

The fragment with gentamicin gene was completely sequenced using primer walking method. The total 4288 bp fragment consists of 2491 bp of foreign DNA and 1797 bp vector backbone (Fig. 3). The BLAST of foreign DNA indicated that the insert were from *L. plantarum* chromosome DNA (2131903-2134583). It contains the entire 1317 bp (2131984-2133300) *murA2* (UDP-N-acetylglucosamine 1-carboxyvinyltransferase) and part of 120 bp (2134583-2134702) *atpC* (H(+)-transporting two-sector ATPase, epsilon subunit). Between *murA2* and *atpC* there is a 973 bp (2133301-2134273) gap with unknown function. There is also an additional 81 bp (2131903-2131983) after *murA2* followed by bla promoter and the Gentamycin resistance gene included in the vector used for genomic integration.



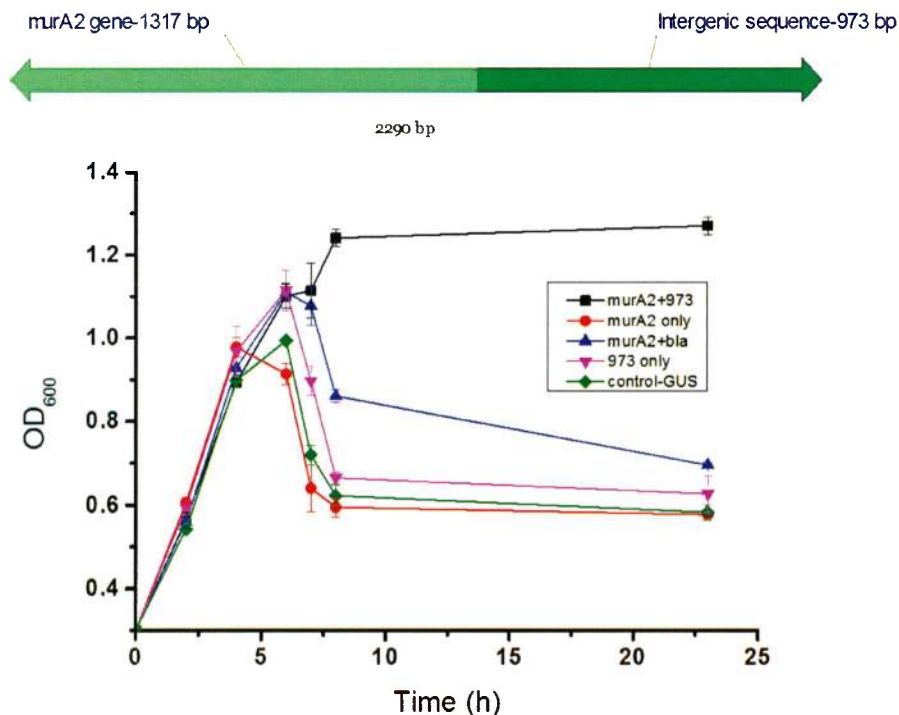
**Fig. 3:** Diagram of sequenced fragment.

**Characterization of sequenced gene fragment.** In order to examine the functionality of the identified gene fragment, the whole fragment including *murA2*, *atpC*, and the additional intergenic region was cloned into a plasmid and expressed in an *E. coli* host to test ethanol tolerance. The entire insert (2491 bp) and the plasmid pDEST-murA2+973 containing the combination (2290 bp) of *murA2* and the intergenic 973 bp region were constructed and transformed into EC100 strain. The control plasmid (pDEST-gus) carries the promoterless *gus* gene coding a beta-glucuronidase from *Arabidopsis thaliana* was constructed as well and served as negative control. The strain EC100gmLP3 was utilized as positive control. Cultivations were carried out in LB medium with 5% (V/V) ethanol at 30 °C. Compared to the negative control, significantly increased maximum density was observed in strains expressing the entire insert (2491 bp), the combination (2290 bp) and the positive control (Fig. 4). The strain harboring only the *murA2* gene with the 973 bp intergenic region reached the highest densities during the 48 hour exposure outperforming both the entire 2491 bp insert and both controls.



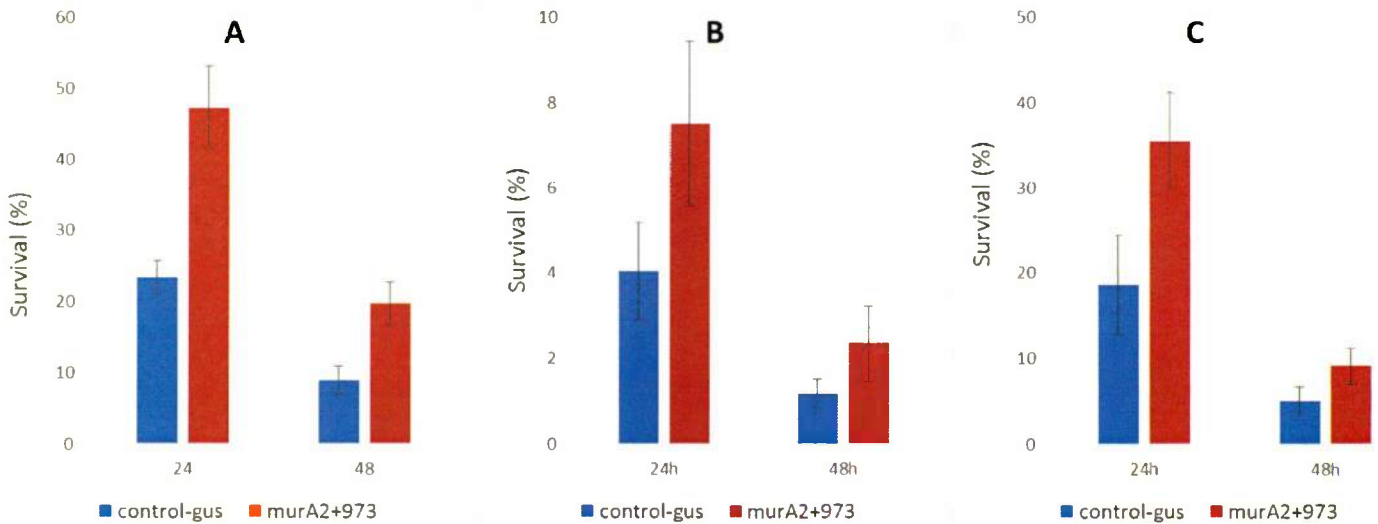
**Fig. 4:** Growth comparison of EC100gmlp3, the strain harboring entire 2491 bp insert and the strain harboring the murA2 gene and its upstream 973 bp sequence. Cells were grown in LB medium with 5% (V/V) ethanol at 30°C and initial optical density of 0.1.

**Plasmid based ethanol tolerance test in EC100 strain.** To further investigate the murA2 gene and its upstream 973 bp region, three plasmids, pDEST-murA2 carrying murA2 gene only, pDEST-973 carrying 973 bp region only and pDEST-murA2-bla carrying murA2 gene under the control of a constitutive *bla* promoter, were constructed and transformed into EC100 strain. Three strains harboring pDEST-murA2, pDEST-973 and pDEST-murA2+bla were cultured in LB medium with 5% (V/V) ethanol and grown at 37 °C. The strain containing pDEST-murA2+973 and the strain harboring pDEST-gus were served as positive control and negative control, respectively (Fig. 5). The improved growth phenotype was observed only when both murA2 and the 973 bp were expressed. Individual expression of either fragment did not improve growth. Expression of murA2 under regulation of the bla promoter improved growth, but had a decreased effect compared to expression of murA2 under its native promoter with the 973 bp intergenic region.



**Fig. 5:** Growth comparison of EC100 strains harboring plasmids containing various fragments from *L plantarum*. Cells were grown in LB medium with 5% (V/V) ethanol and starting optical density of 0.3.

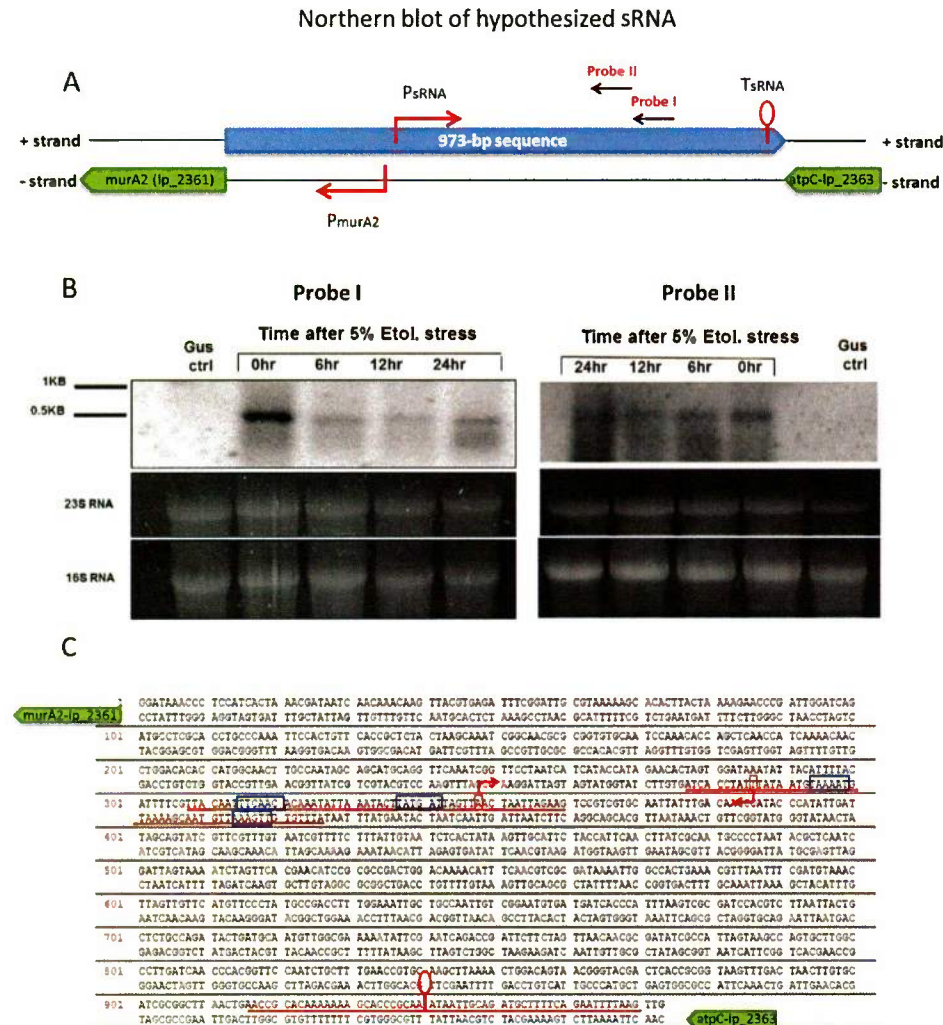
**Survival comparison in *E. coli* K-12 MG1655 with ethanol, n-butanol and iso-butanol stress.** In order to test the effect of the murA2 gene and 973 bp intergenic region on tolerance in wild type *E. coli* the plasmid containing murA2 and the 973 bp intergenic region was transformed into MG1655 cells. The survival of that strain in ethanol, n-butanol, and i-butanol was then compared to a plasmid control (control-gus) (Fig. 6). At 24h, the strains containing murA2+973 obtained 2-fold increase in survival rate compared to the control strain (Fig. 6A) in ethanol. Survival in 1% (v/v) n-butanol and 1% (v/v) iso-butanol was increased by 1.8-fold and 1.9-fold respectively after 24 hours exposure



**Fig. 6:** Comparison of the average survival rates of the MGI655 strains harboring plasmids containing the combination of murA2 gene and 973 bp intergenic sequence (pDEST-murA2+973) and the control plasmid (pDEST-gus) with 3 different alcohols. (A) Cells were grown in LB medium with 50 g/l ethanol. (B) Cells were grown in LB medium with 1% (V/V) n-butanol. (C) Cells were grown in LB medium with 1% (V/V) iso-butanol.

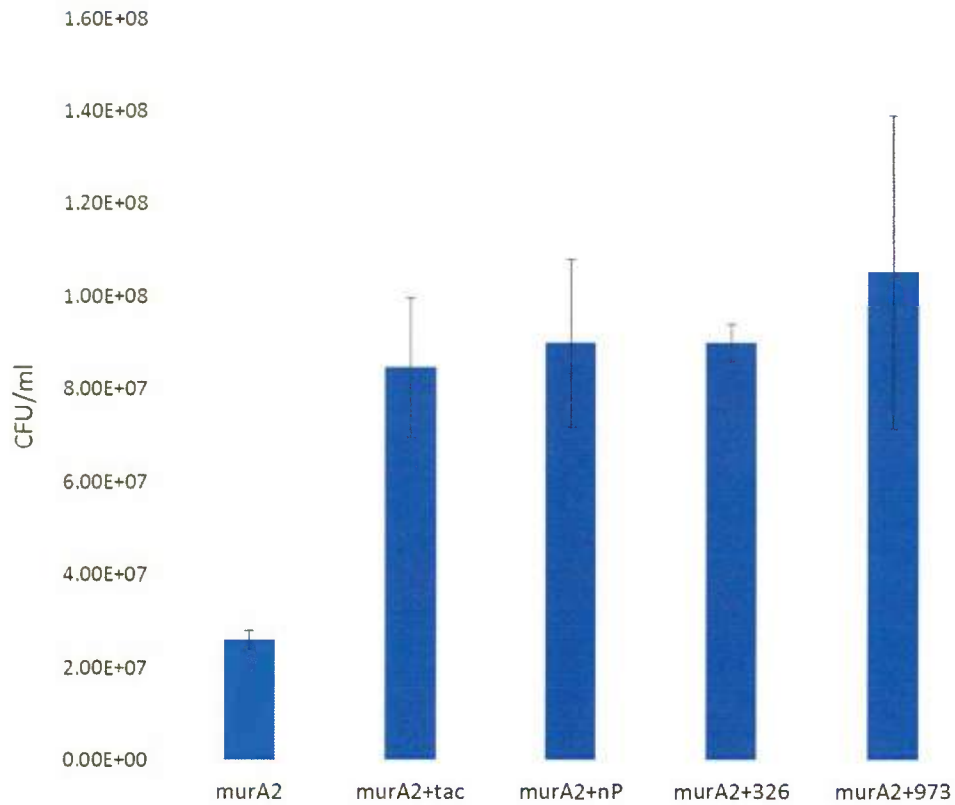
**Characterization of upstream 973-bp intergenic region.** Promoter prediction software (<http://nucleix.mbu.iisc.ernet.in/prombase/>) and ([http://www.fruitfly.org/seq\\_tools/promoter.html](http://www.fruitfly.org/seq_tools/promoter.html)) identified a 326 bp upstream region encoding a promoter for murA2 and a long untranslated region (UTR) (276 bp) prior to the ribosome binding site (Fig. 7C). Additionally, a promoter transcribing in the opposite direction was also predicted in the 973 bp region. Based upon the aforementioned data we developed a theory that the upstream 973 bp intergenic region was producing a regulatory small RNA. Northern blot analysis was performed on the intergenic region to determine if transcription was occurring (Fig. 7). The strain harboring pDEST-gus was served as negative control. Clear single bands shown in each sample lane with the size of 500-700 nt (Fig. 7B) by using two independent probes: (Probe I: cattgcatcagtatctggcagagcag. Probe II: agcagtaattaagacgtggatcgcca). To verify the predicted promoter on the negative strand as well as identify the functionality of the transcribed RNA on positive strand between the predicted promoter and terminator the plasmid pDEST-murA2+326 containing the murA2 gene and its upstream 326 bp (covered exactly the predicted promoter), the plasmid pDEST-murA2+tac containing the murA2 gene under the control of a strong tac promoter and a plasmid Pdest-murA2+nP containing the murA2 gene controlling by its native promoter and RBS (without the 276 bp 5' UTR) were constructed and transformed into *E. coli*. These three strains together with the negative control (pDEST-murA2) and the positive control (pDEST-murA2+973) were cultivated in LB medium in the presence of 5% (V/V) ethanol for 24h. colony forming units (CFU) per ml were obtained and shown in Fig. 8. Expression of murA2 on the strong tac promoter, expression under its native

promoter without the 276 bp UTR, and expression with the full 326 bp UTR produced no statistically significant change from expression of murA2 with the entire 973 bp intergenic region. Expression of murA2 alone however had considerably lower (by 4-fold) viable cell counts compared to expression of murA2 with the 973 bp intergenic region.



**Fig. 7:** Northern blotting analysis of transcribed RNA from 973 bp intergenic region. A: Two single-strand oligonucleotide probes were designed for the potential transcription on the positive strand produced from the region between predicted promoter ( $P_{sRNA}$ ) and terminator ( $T_{sRNA}$ ). Both orientations were indicated by arrows. The arrow on the negative strand indicated the location of predicted promoter ( $P_{murA2}$ ) controlling murA2 gene. B: Transcription confirmation using Northern blot analysis. EC100 cells harboring murA2 gene and its upstream 973 bp were grown in LB medium with 5% (v/v) ethanol. The control cells harboring the plasmids with gus gene. Samples were taken from time points of 0h (overnight seeds culture), 6h, 12h and 24h. The sample of control was taken from 6h. C: DNA sequence of the entire 973 bp intergenic region upstream of murA2 gene. Arrows indicated the

orientations of predicted promoters. (ϕ) indicated the position of predicted terminator. Both predicted promoters and terminator were underlined.



**Fig. 8:** Comparison of ethanol tolerance over CFU in EC100 strain. Cells were grown in LB medium with 5% (v/v) ethanol and initial optical density of 0.2. Samples were taken at 24 h and diluted appropriately before plating on LB agar plates supplied with 100 ug/ml Ampicillin.

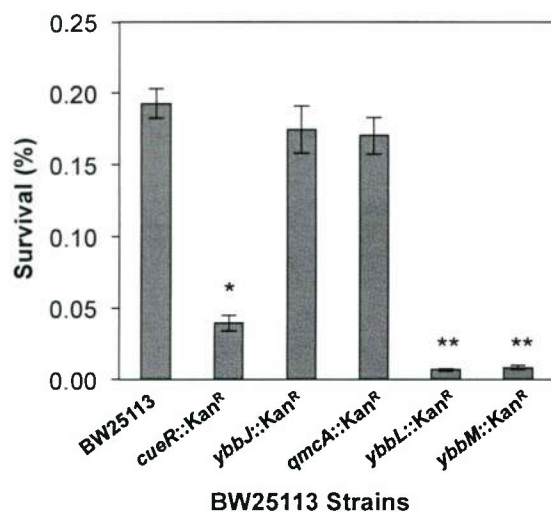
### **Objective 1a:**

#### **Screen genomic libraries for resistance to oxidative stress**

Owing to their high toxicity and associated DNA damage, screening for genomic fragments that enhance resistance to ROS is important for metabolic engineering/bioprocessing applications. An *E. coli* plasmid library was screened for H<sub>2</sub>O<sub>2</sub> tolerance by growth to an OD<sub>600</sub> of 1, and then utilizing a 2% inoculum in media with 14 mM H<sub>2</sub>O<sub>2</sub> that was stressed for 30 min. Dilutions of the surviving cells were plated, and surviving clones were further characterized. The selection protocol utilized stress with a high H<sub>2</sub>O<sub>2</sub> concentration of 14 mM over 30 minutes to ensure that most cells would be killed, therefore eliminating false positives. Plasmids from strains with higher H<sub>2</sub>O<sub>2</sub> survival were isolated and sequenced. Analysis identified two inserts spanning the same chromosomal region in opposite

orientations. One plasmid contained a ~3.8 kb insert (chromosomal location 513,453 – 517,242) with an incomplete *cueR* and *ybbN* and the complete ORFs of *ybbJ*, *qmcA*, *ybbL* and *ybbM*. The second plasmid contained a ~4.3 kb insert (chromosomal location 513,076 – 517,376) spanning the same chromosomal region in the opposite orientation with the addition of a complete *cueR* ORF. The larger plasmid was designated **pLib** and was further characterized in assays with even higher oxidative stress.

To investigate the effect of the genes included in the pLib plasmid on H<sub>2</sub>O<sub>2</sub> tolerance the knockouts (KOs) of these genes created in *E. coli* BW25113 as part of the Keio Collection were tested (Baba et al., 2006). The survival rate of the individual gene KOs after 30 minute stress with 4 mM H<sub>2</sub>O<sub>2</sub> at 37 °C in LB was determined (Fig. 9). The  $\Delta ybbJ$  and  $\Delta qmcA$  strains exhibit similar survival to the parent strain, indicating that they have no direct role in ROS response. The disruption of the copper regulatory gene *cueR* decreases tolerance to H<sub>2</sub>O<sub>2</sub> substantially. The  $\Delta ybbL$  and  $\Delta ybbM$  KOs also exhibit more sensitivity to H<sub>2</sub>O<sub>2</sub> and are even less tolerant than the  $\Delta cueR$  strain (Fig. 9). *ybbL* and *ybbM* encode for predicted ATP-binding component and the inner membrane of an unknown ABC transporter that might have a role in metal resistance. Although the role of these genes was not known, their disruption significantly decreased tolerance to oxidative stress.

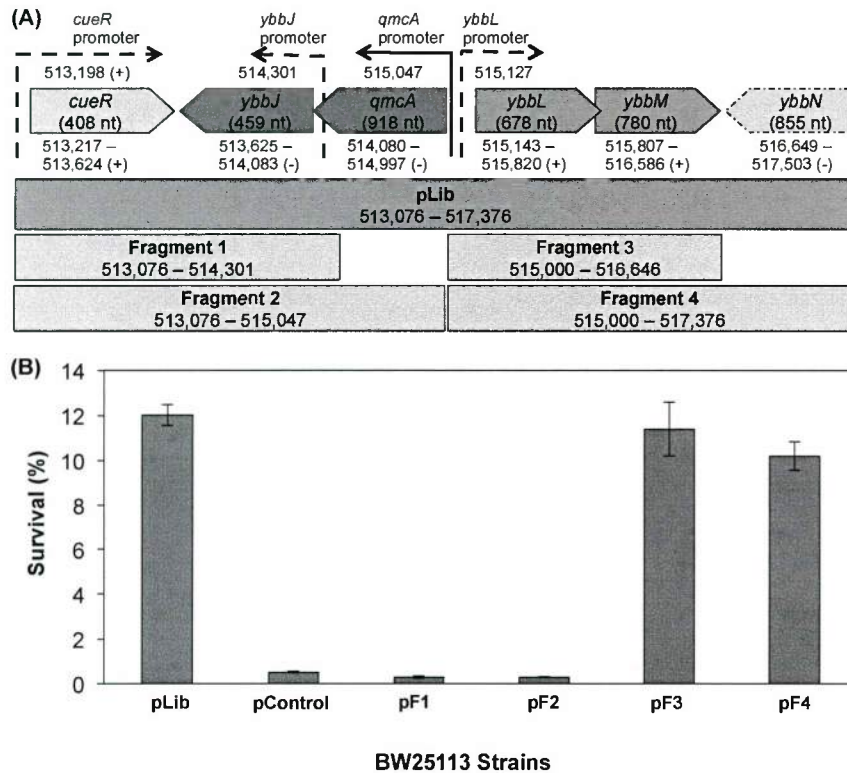


**Fig. 9:** Survival rates of *E. coli* BW25113 and the *cueR*::Kan, *ybbJ*::Kan, *qmcA*::Kan, *ybbL*::Kan and *ybbM*::Kan strains of the Keio Collection (Baba et al., 2006) after a 30 minute exposure to 4 mM H<sub>2</sub>O<sub>2</sub> at 37 °C. N = 3 mean ± SEM. The  $\Delta cueR$  has lower survival compared to the BW25113 strain, as determined by a t-test (p-value < 0.05) and indicated by the (\*). The *ybbL*::kan and *ybbM*::kan strains had lower survival than the *cueR*::kan strain (\*\*) (p-value = 0.014 and 0.016, respectively).

### Multicopy expression of the *ybbL-ybbM* operon confers resistance to oxidative stress

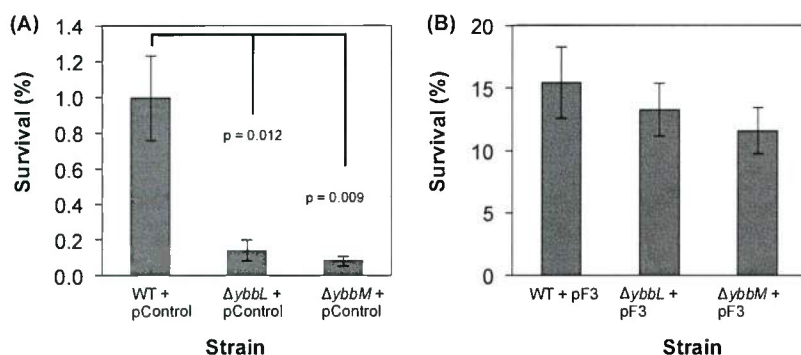
To identify which gene included in pLib conferred the higher resistance to oxidative stress; fragments of the genomic region were generated and individually tested. We created four different fragments that span pLib (Fig 10A). The constructs were transformed in *E. coli* BW25113 and the effect of each fragment was investigated in H<sub>2</sub>O<sub>2</sub> assays (Fig. 10B). The original pLib plasmid was also used, as well as a plasmid control containing the promoterless *Arabidopsis thaliana* beta-glucuronidases (*gus*) gene to maintain the same antibiotic stress and metabolic burden of plasmid propagation on all strains. As demonstrated in Fig. 10B, fragments 1 and 2 have no effect on the H<sub>2</sub>O<sub>2</sub> tolerance of *E. coli* BW25113. Fragments 3 and 4 have statistically equal survival to the strain harboring the pLib. The overexpression data presented here is consistent with the knockout data, and thus we conclude that the *ybbL* -

*ybbM* operon is responsible for the higher tolerance observed.



**Fig. 10:** Determining which pLib genes have a beneficial effect on H<sub>2</sub>O<sub>2</sub> tolerance. **(A)** The ORFs of the genes included in the pLib plasmid that enhances H<sub>2</sub>O<sub>2</sub> tolerance and the different fragments constructed. The pLib plasmid contains the ORFs of *cueR*, *ybbJ*, *qmcA*, *ybbL* and *ybbM* (solid lines). The fragment also contains an incomplete *ybbN* (dashed lines). 4 fragments were created: Fragment 1 contains *cueR* & *ybbJ*, Fragment 2 contains *cueR*, *ybbJ* and *qmcA*, Fragment 3 contains *ybbL* & *ybbM*, and Fragment 4 contains *ybbL*, *ybbM* and the fragment of *ybbM* contained in pLib. **(B)** H<sub>2</sub>O<sub>2</sub> tolerance of different fragments. Survival rates of *E. coli* BW25113 with plasmids containing different fragments after a 30 minute exposure to 4 mM H<sub>2</sub>O<sub>2</sub> at 37 °C. The pLib is the plasmid isolated during library enrichment. pControl contains the promoterless *Arabidopsis thaliana* beta-glucuronidases (*gus*). pF1, pF2, pF3 and pF4 correspond to fragments 1 – 4 illustrated in (A). N = 4 mean ± SEM.

To validate our overexpression in pF3, we performed quantitative reverse-transcription PCR on BW25113 strains harboring a plasmid control or pF3, and determined that the expression of *ybbL* and *ybbM* are increased by  $49 \pm 8$  and  $35 \pm 10$ , respectively, relative to the housekeeping gene *ihfB*. Furthermore, pF3 was transformed into the *ybbL*::kan and *ybbM*::kan strains to determine if the phenotype can be restored. Fig. 12 illustrates that pF3 recovers the phenotype in the *ybbL* and *ybbM* KO strains, and increases tolerance past the WT with a plasmid control (Fig. 11A) to the levels of the WT harboring the beneficial pF3 (Fig. 11B).

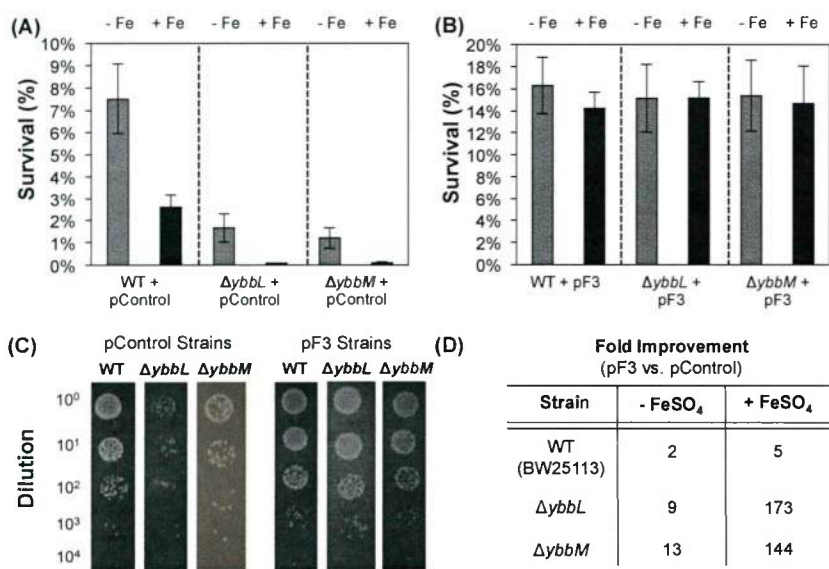


**Fig. 11:** Reconstitution of *ybbL* and *ybbM* (expressed from plasmid pF3) in the gene knockout strains recovers the phenotype and improves tolerance to H<sub>2</sub>O<sub>2</sub> stress. Survival rates of *E. coli* BW25113 (WT), *ybbL*::Kan and *ybbM*::Kan strains of the Keio Collection (Baba et al., 2006) with a control plasmid (A) or the beneficial pF3 (B) after a 30 minute exposure to 4 mM H<sub>2</sub>O<sub>2</sub> at 37 °C in LB media. N = 5 mean  $\pm$  SEM. P-values indicate the statistical significance of a t-test performed between the WT and the KO strains.

#### The function of YbbL and YbbM: a metal ABC transporter associated with iron

YbbL and YbbM have not been previously characterized but belong to the ABC family of transporters, with YbbL being the predicted ATP binding component subunit and the YbbM being the inner membrane protein that allows for transport. However, the metal substrate of this ABC transporter could not be inferred bioinformatically. We hypothesized that the ABC transporter encoded by *ybbL* and *ybbM* enhances resistance to H<sub>2</sub>O<sub>2</sub> via the intracellular export of a metal that can partake in Fenton reactions directly or indirectly by eliminating iron from active sites. By eliminating the free metal, fewer oxidative reactions can take place in the presence of H<sub>2</sub>O<sub>2</sub> and tolerance is enhanced.

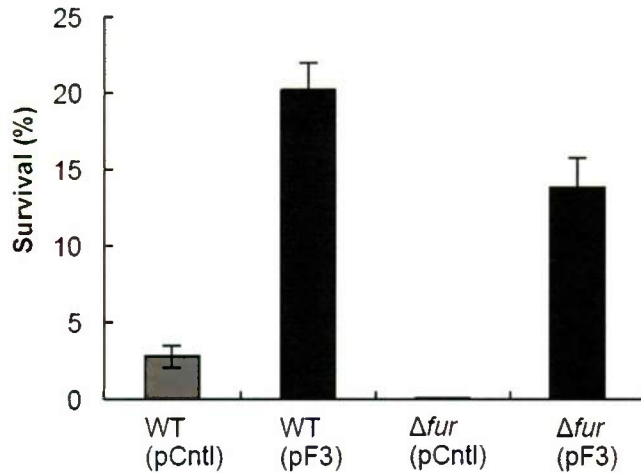
To determine the metal substrate, we investigated strains overexpressing the YbbL-YbbM transporter in M9 minimal media supplemented with metals. ABC transporters can have an effect on multiple metals; thus, we decided to investigate the role of Al<sub>2</sub>(SO<sub>4</sub>)<sub>3</sub> (100  $\mu$ M), CuSO<sub>4</sub> (20  $\mu$ M), NiCl<sub>2</sub> (30  $\mu$ M), and FeSO<sub>4</sub> (30  $\mu$ M). We performed H<sub>2</sub>O<sub>2</sub> survival assays on the BW25113 *ybbM*::Kan strain and observed a profound decrease in survival when iron was present, but no major effect with other metals. We proceeded to investigate the H<sub>2</sub>O<sub>2</sub> tolerance of all strains when grown in M9 minimal media with 30  $\mu$ M FeSO<sub>4</sub> (Fig. 12). Iron is inhibitory to all strains harboring the plasmid control (Fig. 12A), and decimates the  $\Delta ybbL$  and  $\Delta ybbM$  strains, resulting in survival rates of 0.14% and 0.08%, respectively. However, the presence of FeSO<sub>4</sub> does not have a discernible effect on survival of the strains harboring the overexpression pF3 plasmid (Fig. 12B). This demonstrates that this ABC transporter must have a role in iron transport, and by maintaining optimal cellular metal levels tolerance to H<sub>2</sub>O<sub>2</sub> is increased.



**Fig. 12:** Survival under H<sub>2</sub>O<sub>2</sub> stress in M9 minimal media containing 30  $\mu$ M FeSO<sub>4</sub>. **(A)** *E. coli* BW25113 (WT), *ybbL::Kan* and *ybbM::Kan* strains of the Keio Collection (Baba et al., 2006) with a control plasmid were grown in minimal media in the presence (black bars) or absence (grey bars) of 30  $\mu$ M FeSO<sub>4</sub>. The cultures were then stressed with 4 mM H<sub>2</sub>O<sub>2</sub> at 37 °C in M9 minimal media with 30  $\mu$ M FeSO<sub>4</sub> for 30 mins and survival rates were determined based on colony forming units. The survival of the WT strain is higher than the KO strains both in the presence and absence of iron. **(B)** The strains with the pF3 overexpression plasmid were treated similarly to (A). The survival of all overexpression strains is statistically indistinguishable. N = 3 mean  $\pm$  SEM. **(C)** Final dilution series of assays performed in M9 media with 30  $\mu$ M FeSO<sub>4</sub> are shown for the pControl strains (left) and the overexpression pF3 strains (right). For these assays 10  $\mu$ l of diluted culture post stress was spotted on M9 agar plates for visual representation. **(D)** The fold improvement of the different strains harboring the overexpression pF3 or the pControl plasmids was determined for the assays in the presence or absence of FeSO<sub>4</sub>.

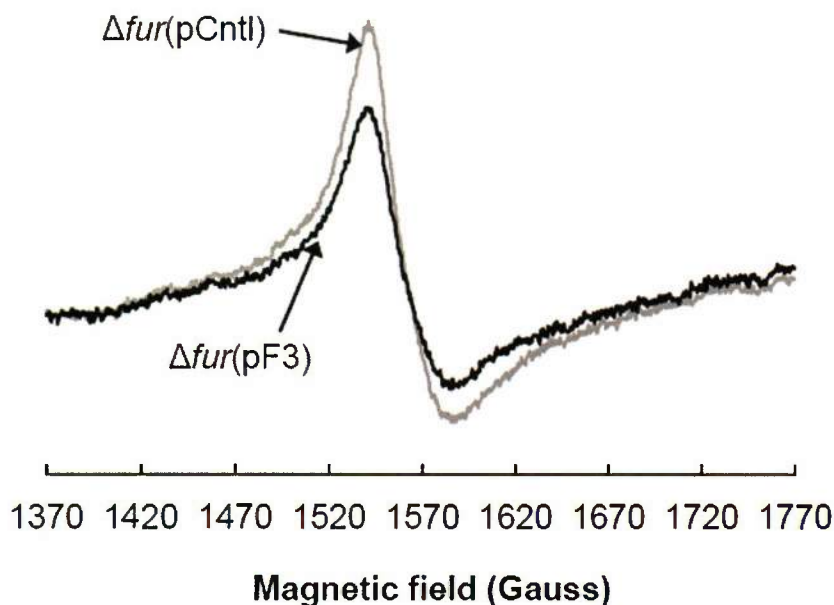
***fetA* and *fetB* overexpression guards against lethal oxidative damage in  $\Delta fur$  mutants.** The Fur protein regulates iron uptake and superoxide dismutase expression (Andrews et al., 2003), with *fur* mutants being unable to balance cellular iron content, thus resulting in iron overload (Touati et al., 1995). Iron accumulation in *fur* mutants results in severe oxidative damage and mutagenesis, with *fur recA* mutants being unable to survive under aerobic conditions (Touati et al., 1995). We next investigated if overexpressing *fetA* and *fetB* can overcome the deficiencies of the  $\Delta fur$  strain. Survival of the  $\Delta fur$ (pF3) strain increased in the presence of 30  $\mu$ M FeSO<sub>4</sub> (Fig. 13), demonstrating that FetA and FetB have a role in overcoming the iron overload elicited by the *fur* mutation. This result corroborates the previous data (Fig. 12) where we demonstrate that *fetA* and *fetB* overexpression can abrogate the negative effects of iron overload coupled with H<sub>2</sub>O<sub>2</sub> stress. Furthermore, this finding suggests that the FetAB transporter is not regulated by Fur directly. The Fur consensus

sequence is not found upstream of *fetA* or *fetB* (Lavrrar et al., 2002), suggesting that the transporter is regulated by other means.



**Fig. 13:** *E. coli* BW25113 WT and  $\Delta fur$  mutant survival under  $H_2O_2$  stress in LB media containing  $30 \mu M FeSO_4$ . The cultures were grown in media containing  $30 \mu M FeSO_4$  to facilitate iron overload. Data are means  $\pm$  SEM ( $n = 5$ ).

**Overexpression of *fetA* and *fetB* decreases intracellular iron in  $\Delta fur$ .** The iron content in BW25113  $\Delta fur$  was investigated by electron paramagnetic resonance (EPR). Samples prepared from cultures grown aerobically in LB media to exponential phase were prepared and analyzed by EPR. Deregulation of Fur results in iron overload. We found that the total intracellular free iron content of  $\Delta fur$  with *fetA* and *fetB* overexpression was lower than in the control (Fig. 6). The peak-to-peak amplitude of five biological samples for  $\Delta fur(pF3)$  and  $\Delta fur(pCntl)$  were compared by a t-test ( $p$ -value of 0.030) and the amplitude was significantly decreased by 37%. The  $\Delta fur$  stains are still overloaded with iron, but overexpression of *fetA* and *fetB* resulted in lower cellular iron content. Because ABC transporters operate between the cytoplasm and the periplasm, FetAB would export iron from the cytoplasm to the periplasm, where it can still be detected by whole-cell EPR. The overall iron content of the cells would not be expected then to decrease dramatically. This would explain why the amplitude of the signals for the  $\Delta fur(pF3)$  strain are high even with *fetA* and *fetB* overexpression (Fig. 14). These EPR iron data combined with the  $H_2O_2$  stress data (Fig. 14) strongly suggest that FetAB is involved in iron export.



**Fig. 14:** Overexpression of *fetA* and *fetB* in *E. coli* BW25113  $\Delta fur$  decreases the overall intracellular free iron concentration as determined by electron paramagnetic resonance (EPR). The overexpression strain BW25113  $\Delta fur(pF3)$  is shown in black and the control  $\Delta fur(pCntl)$  in gray. Cultures grown aerobically in LB media in exponential phase (OD600 of 1 – 2) were harvested, and were oxidized with desferrioxamine mesylate (DF) as described in (Woodmansee & Imlay, 2002). Samples loaded into 4 mm thin wall precision Quartz EPR sample tubes (Wilmad) were analyzed on an a Bruker EMXmicro X-band spectrometer with variable temperature unit (ER4141VT) with measurement settings as follows: 30.00 milliwatts; microwave frequency, 9.37 GHz; modulation frequency, 100 kHz; modulation amplitude, 10.00 G; sweep rate, 4.88 G/s; time constant, 81.92 ms. The signals were normalized by protein concentration. Five biological replicates were analyzed and a representative sample is shown.

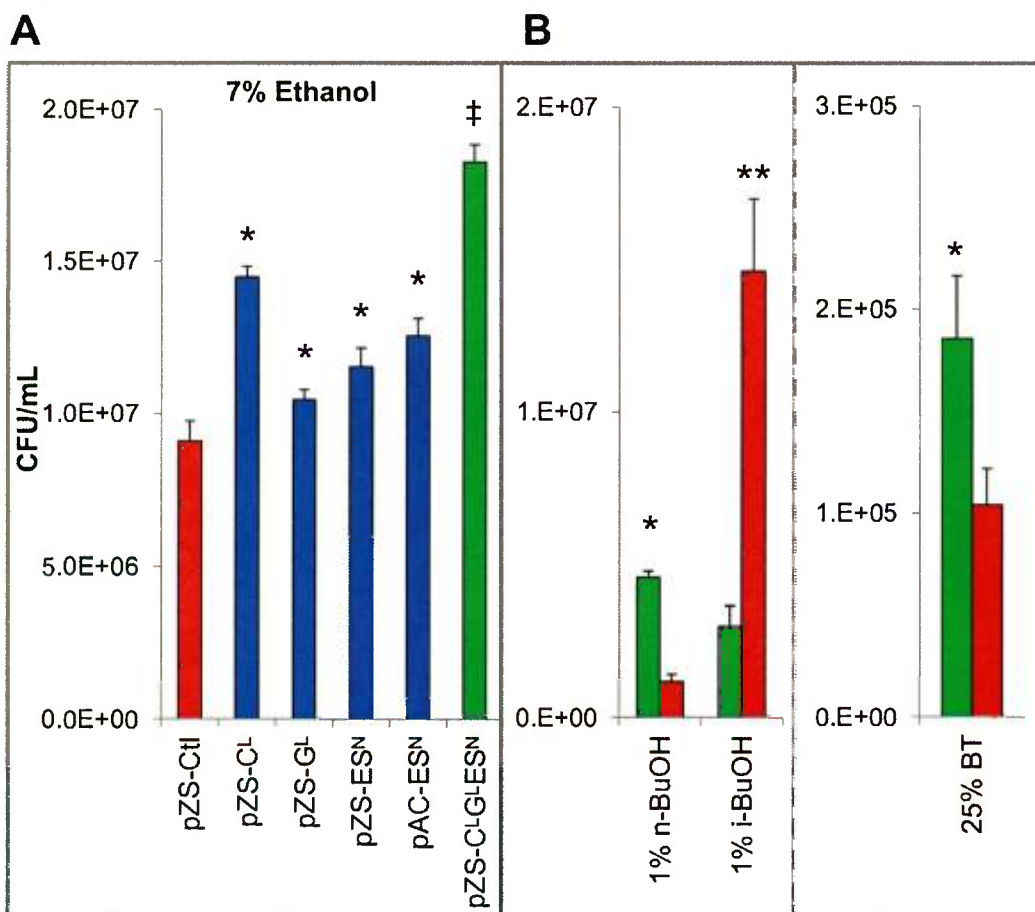
## Objective 2

**Overview.** Our initial work in the development of a customizable stress-response system began with the *E. coli* genes *groES* and *groEL*. Utilizing their native promoter and a plasmid based expression platform we demonstrated a significant improvement to ethanol and BT tolerance. We then focused on optimizing the expression of the four HSP genes (*groES*, *groEL*, *grpE*, *clpB*), alone and in combination for improved tolerance to ethanol, BT, and other solvents. These genes were then tested for improved ethanol productivity from the ethanologenic *E. coli* KO11 strain in combination with the *murA2* fragment identified through screening of *L. plantarum* genes integrated into the *E. coli* genome described above. Finally, we examined methods for increasing the pool of genes in our stress-response system

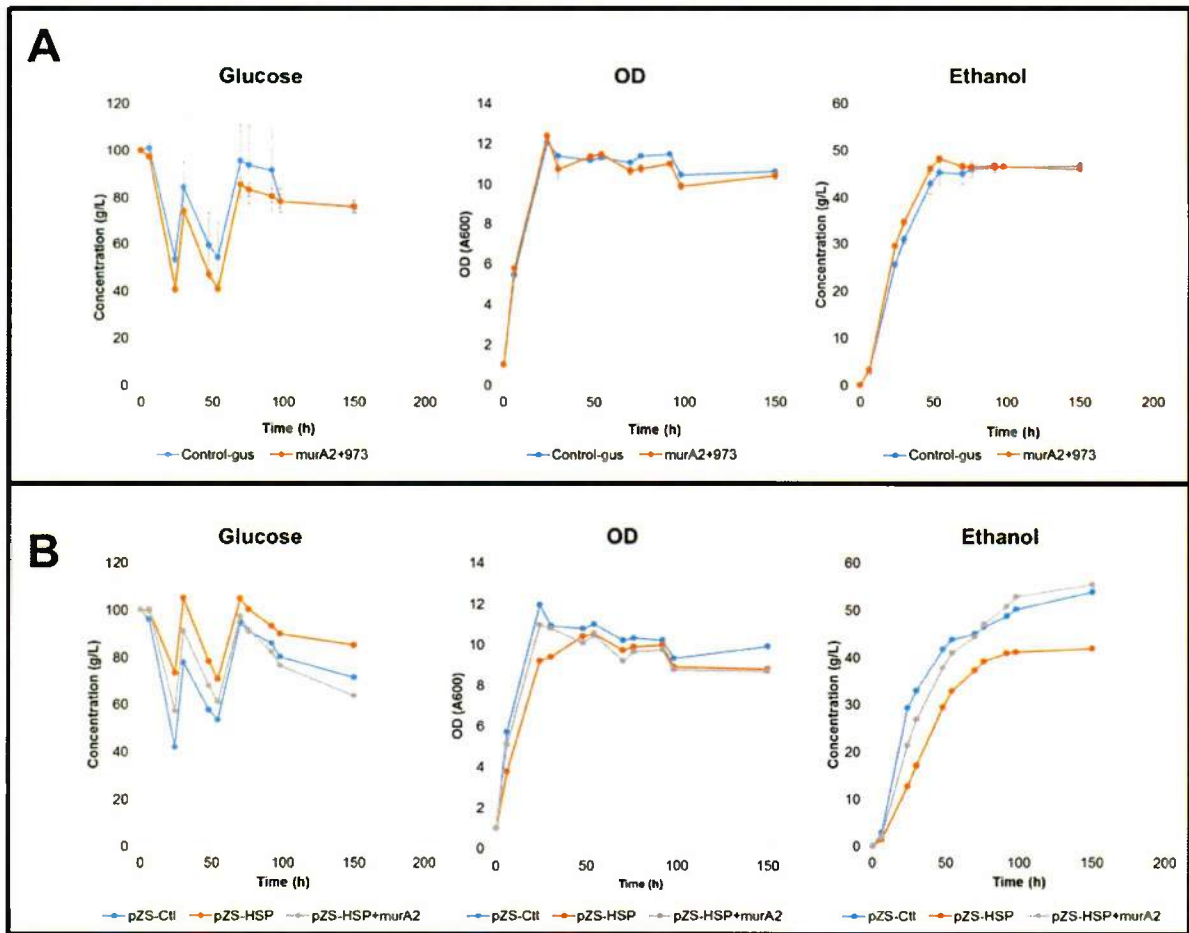
beyond the limitations of molecular chaperones using library based screening of heterologous genomes for improved tolerance to ethanol.

**Improving ethanol tolerance using a four-chaperone overexpression system on a single plasmid.** Previously we had constructed the plasmid pZS-C<sup>L</sup>G<sup>L</sup>ES<sup>N</sup> to express the four identified HSPs, GroES, GroEL, ClpB, and GrpE on a single expression vector. The resulting strain was compared to overexpression of each of the individual genes using pZS-C<sup>L</sup>, pZS-G<sup>L</sup>, pZS-ES<sup>N</sup>, and pAC-groESL as well as the plasmid control pZS-Ctl for tolerance to 7% ethanol (Fig. 15). The four-gene overexpression strain produced a statistically significant improvement in the ratio of CFUs from 0 to 24 hours compared to all of the other strains and a 2.5-fold increase in the ratio from the control plasmid. This improvement indicates that the chaperones expressed are acting beneficially together and provides the basis for general stress response system, which can be tuned to a particular solvent. Additionally, the effects of the HSP plasmid were compared to the plasmid control for tolerance to 1% n-butanol, 1% i-butanol, and 25% BT. In n-butanol and BT overexpression of the HSPs produced a 3.9-fold and 78% respective increase in viable cell counts compared to the control. However, expression of the HSPs in i-butanol caused an 80% lower viable cell count than that of the control. This less effective response to i-butanol stress suggests that this combination of genes under these expression conditions may be suboptimal for handling i-butanol stress.

**Testing chaperone overexpression system on production of ethanol in *E. coli*.** To evaluate the effects of the observed improved ethanol tolerance generated both by expression the murA2 gene from *L. plantarum* and overexpression of the HSPs identified we examined the impact on ethanol production using the *E. coli* KO11 strain (Ohta et al., 1991). Glucose consumption, cell density, and ethanol production are shown the strains examined (Fig. 16). Expression of the *L. plantarum* murA2 fragment appears to produce a slight increase in production rate of ethanol during the first 48 hours compared to its plasmid control. However, this effect is not statistically significant and no increase in final titer is observed. Expression of the HSP plasmid appears to slow both growth and production throughout the entire fermentation. However, this may be an effect of overexpression of the HSPs during the early portions of the fermentation where there are low concentrations of ethanol and therefore less need for their expression. Overexpression of the HSPs at the beginning of the fermentation may therefore be undesirable and act as a drain on other cellular processes including growth and production. Using a more controlled expression system it may be possible to induce expression of the HSPs only during the late stages of the fermentation where ethanol concentrations have reached more toxic levels. Notably, the combination of overexpression of the HSPs and the murA2 fragment on a single plasmid restores the growth phenotype of the control and appears to possibly increase productivity during the late stages of the fermentation. This may be due to possible beneficial interactions between MurA2 and the overexpressed *E. coli* HSPs. We are currently examining alternative growth conditions including larger bioreactor fermentations and later induction to attempt to improve productivity in our engineered strains.



**Fig. 15:** (A) Viable cell counts (CFU/mL) performed after 24 hours of culture in 7 % ethanol for the 4-gene, single plasmid overexpression system and 0.1 mM IPTG induction. (B) CFU/mL performed after 24 hours of culture in 1% n-butanol, 1% i-butanol, and 25% 1,2,4-butanetriol (BT). MG1655 (pZS-Cl<sup>L</sup>G<sup>L</sup>ES<sup>N</sup>) (Green); MG1655 (pZS-Ctl) (Red). Results are from 3 biological replicates; error bars indicate standard error between replicates. \* - Indicates statistically significant improvement from control strain, MG1655 (pZS-Ctl) ( $p < 0.05$ ). \*\* Indicates statistically significant increase above overexpression strain MG1655(pZS-Cl<sup>L</sup>G<sup>L</sup>ES<sup>N</sup>). ‡ - Indicates statistically significant increase from all other strains examined ( $p < 0.05$ ).

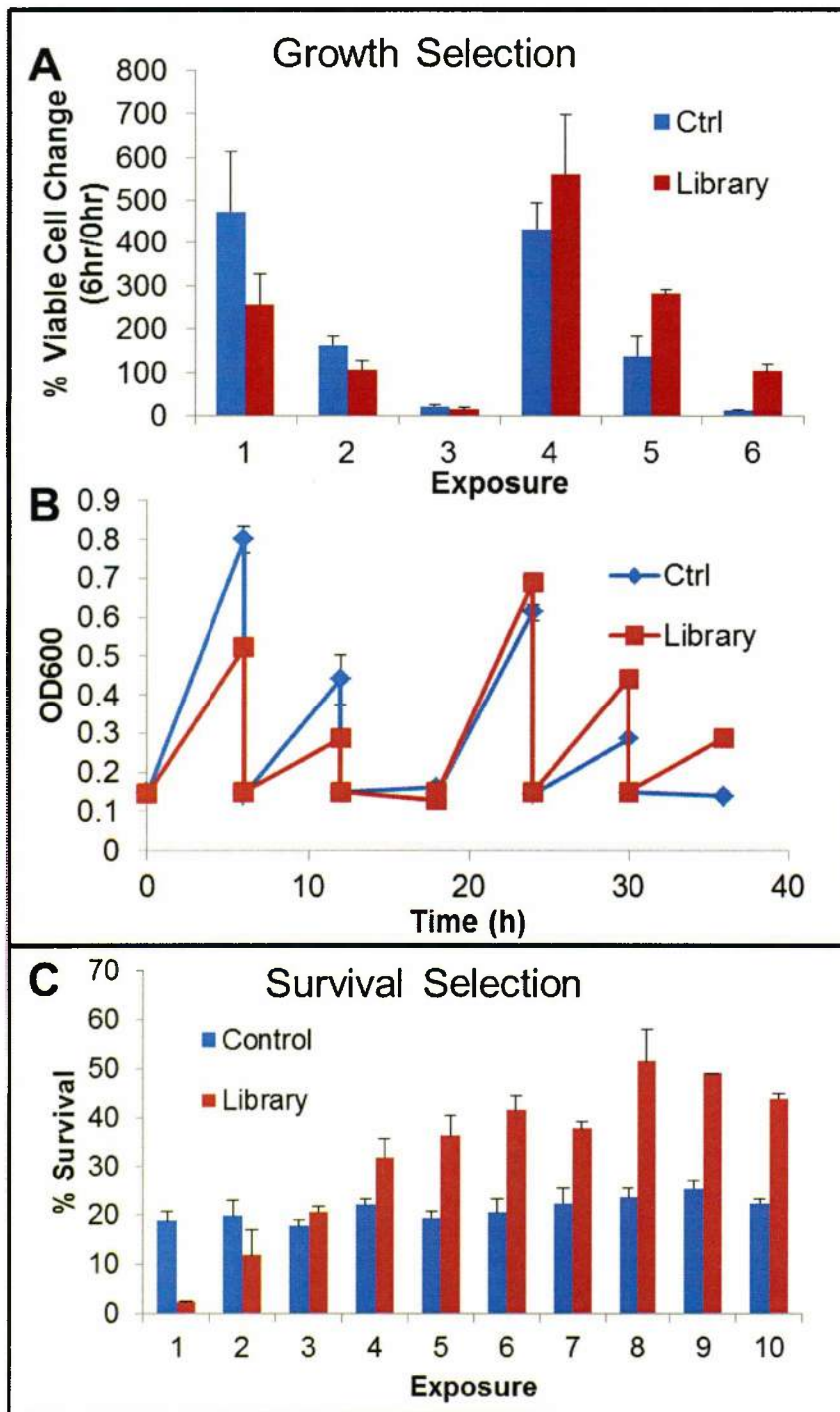


**Fig. 16:** Flask based ethanol fermentations using the *E. coli* KO11 strain bearing tolerance enhancing plasmids. **(A)** Glucose consumption, cell density, and ethanol production of KO11 cells containing a plasmid control (Control –gus) and a plasmid containing the *L. plantarum* murA2 fragment (murA2+973). **(B)** Glucose consumption, cell density, and ethanol production of KO11 cells containing a plasmid control (pZS-Ctl), the HSP plasmid pZS-C<sup>L</sup>G<sup>L</sup>ES<sup>N</sup> (pZS-HSP), and a plasmid containing the *L. plantarum* murA2 fragment (murA2+973) and the HSPs (pZS-HSP+murA2). Data represent biological triplicate data.

**Identifying heterologous genes for use in the stress response system.** In order to further expand the list of gene targets in our stress response system we sought to identify new genes from more tolerant, heterologous prokaryotic species. Using an *E. coli* strain with an integrated, inducible copy of the sigma factor RpoD from *L. plantarum* (MG1655::rpoD) we began screening plasmid based *L. plantarum* genomic libraries in combination with the HSP expression plasmid (pZS-C<sup>L</sup>G<sup>L</sup>ES<sup>N</sup>). The *L. plantarum* sigma factor was included to increase the number of genes expressed in *E. coli* from the heterologous genomic library without relying on external promoters. Identified genomic fragments would thus be expressed, either dependent upon RpoD or on native transcription machinery, and able to function

cooperatively with the overexpressed *E. coli* HSPs. The library constructed contained *L. plantarum* genomic DNA fragments with an average insert size of 4 kb.

In order to isolate beneficial clones for ethanol tolerance two enrichment protocols were used separately. A growth screening using a low concentration, short time exposure assay (4% ethanol for 6 hours) with three exposure cycles followed by a 6 hour recovery step in media containing no stressor and a survival screen using a high concentration, long time exposure (8% ethanol for 24 hours) with a recovery step after every exposure. Both assays used a 0.15 starting optical density for each exposure and 10% inoculation for each recovery step. Percent survival was monitored for the survival assay and cell density and change in viable cell count was monitored for the growth enrichment in biological triplicate (Fig. 17). The enrichment protocol was performed on both the library and a library plasmid control, both also expressing the HSP plasmid pZS- C<sup>L</sup>G<sup>L</sup>ES<sup>N</sup>, to determine if improvements were due to isolation of beneficial library clones or accumulation of random mutations. The growth enrichment presented consistently improved cell density (Fig. 17A) and survival (Fig. 17B) for the library pool during the fourth and all subsequent exposures. The survival selection showed significant improvements for the library over the plasmid control during the fourth and all subsequent exposures (Fig. 17C).



**Fig. 17:** Enrichment results for growth and survival screening of *L. plantarum* plasmid library in *E. coli* MG1655::rpoD cells bearing pZS-C<sup>L</sup>G<sup>L</sup>ES<sup>N</sup>. Strain bearing plasmid control (blue) and *L. plantarum* library (red) were screened in biological triplicate, error bars indicate standard error of the mean. Recovery growth in media without stressant was performed after the third exposure for the growth selection and after each exposure for the survival selection.

(A) Percent change in viable cells over 6 hour exposures during growth screening (4% ethanol, 6 hour exposure). (B) Cell density before and after each exposure during growth screening. (C) Percent survival during survival selection (8% ethanol, 24 hour exposure).

**Table 1:** Sequencing results from *L. plantarum* library growth enrichment.

ID	Start	End	Size	Genes
G1-1	3199864	3203643	3780	(clpL), (lox)
G1-2	3026086	3026192	107	(CcsB family cell surface protein)
G1-3	3199864	3203643	3780	(clpL), (lox)
G1-4	3026086	3026192	107	(CcsB family cell surface protein)
G1-5	3199864	3203643	3780	(clpL), (lox)
G1-6	3199864	3203643	3780	(clpL), (lox)
G1-7	3199864	3203643	3780	(clpL), (lox)
G1-8	2224343	2224446	104	(mucus-binding protein LPXTG-motif cell wall anchor)
G1-9	2224343	2224446	104	(mucus-binding protein LPXTG-motif cell wall anchor)
G1-10	3199864	3203643	3780	(clpL), (lox)
G2-1	1429291	1430754	1464	Lp_1564 membrane protein, (LyTR family transcriptional regulator)
G2-2	1429291	1430754	1464	Lp_1564 membrane protein, (LyTR family transcriptional regulator)
G2-3	13810	17234	3425	(orf19), orf20, orf21, tral, (orf23)
G2-4	1086051	1088519	2469	(rfbC), rfbB, rfbD, (transposase, fragment)
G2-5	1429291	1430754	1464	Lp_1564 membrane protein, (LyTR family transcriptional regulator)
G2-6	1683219	1683417	199	(Lp_1863 drug resistance transport protein, major facilitator subfamily (MFS), EmrB/QacA subfamily)
G2-7	1429291	1430754	1464	Lp_1564 membrane protein, (LyTR family transcriptional regulator)
G2-8	1429291	1430754	1464	Lp_1564 membrane protein, (LyTR family transcriptional regulator)
G2-9	1086051	1088519	2469	(rfbC), rfbB, rfbD, (transposase, fragment)
G2-10	8865	9132	268	(gyrA)
G3-1	3199864	3203643	3780	(clpL), (lox)
G3-2	1086051	1088519	2469	(rfbC), rfbB, rfbD, (transposase, fragment)
G3-3	1024357	1026809	2453	(cydA), cydB, (cydC)
G3-4	3199864	3203643	3780	(clpL), (lox)
G3-5	3199864	3203643	3780	(clpL), (lox)
G3-6	1024357	1026809	2453	(cydA), cydB, (cydC)
G3-7	3150169	3151195	1027	(lp_3533 carbohydrate symporter)
G3-8	3199864	3203643	3780	(clpL), (lox)
G3-9	2299496	2301284	1789	(glpQ4)
G3-10	1086051	1088519	2469	(rfbC), rfbB, rfbD, (transposase, fragment)

**Table 2:** Sequencing results from *L. plantarum* library survival enrichment.

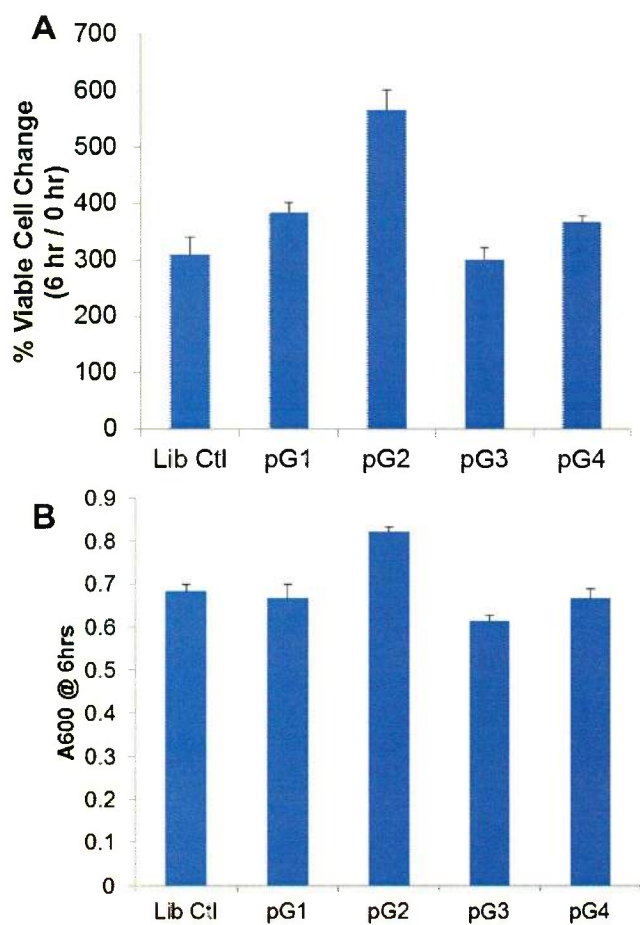
ID	Start	End	Size	Genes in Fragment
S1-1	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S1-2	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S1-3	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S1-4	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S1-5	1004096	1006110	2015	(mae) citC citD (citE)
S1-6	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S1-7	1004096	1006110	2015	(mae) citC citD (citE)
S1-8	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S1-9	74429	77599	3171	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], GNAT acetyltransferase
S1-10	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S2-1	2866824	2868276	1453	(pts26BCA)
S2-2	73338	77406	4069	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S2-3	24728	24860	133	(traE)
S2-4	73338	77406	4069	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S2-5	2866824	2868276	1453	(pts26BCA)
S2-6	3158907	3162902	3996	tal2 [lp_3539], [lp_3540], pts34B, pts34A, (lp_3543)
S2-7	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S2-8	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S2-9	2866824	2868276	1453	(pts26BCA)
S2-10	73338	77406	4069	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S3-1	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S3-2	74983	77022	2040	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078]
S3-3	1004096	1006110	2015	(mae) citC citD (citE)
S3-4	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S3-5	74983	77022	2040	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078]
S3-6	2238662	2239489	828	(hpk7), (rrp7)
S3-7	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)

S3-8	1004096	1006110	2015	(mae) citC citD (citE)
S3-9	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
S3-10	1004096	1006110	2015	(mae) citC citD (citE)

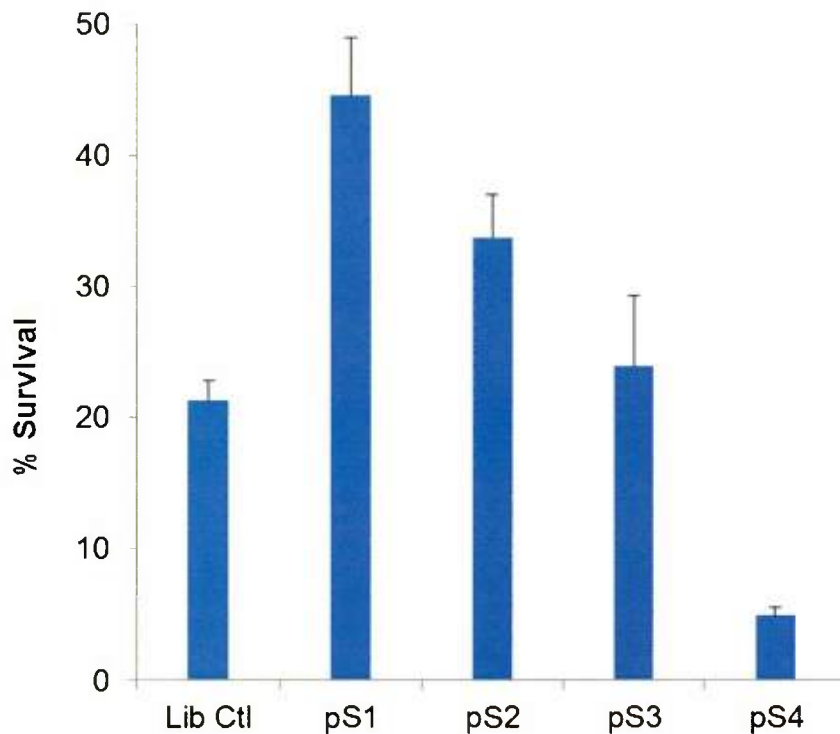
From the library selection replicates 10 colonies were isolated from the sixth exposure in both the growth and survival screening (30 colonies from each enrichment). Library plasmids were isolated and sequenced, the growth sequence results are presented in Table 1 and the survival sequence results are presented in Table 2. From each of these selections the four inserts which were the most prevalent in the pool of sequenced samples for verification of improved ethanol tolerance. The plasmids containing these fragments were renamed (Table 3) and transformed into fresh MG1655::rpoD (pZS- C<sup>L</sup>G<sup>L</sup>ES<sup>N</sup>) to eliminate the impact of possible mutations to either genomic DNA or the HSP plasmid. Tolerance of these strains was then determined using growth and survival assays in ethanol. Cell density and percent change in viable cell counts were monitored for the growth plasmids (pG plasmids) for the growth tolerance assay using 4% ethanol with a 6 hour exposure (Fig. 18). Percent survival was monitored for strains containing the survival plasmids (pS plasmids) in the survival assay using 8% ethanol for a 24 hour exposure (Fig. 19).

**Table 3:** Selected plasmids for verification of improved tolerance phenotype.

Plasmid Name	Start	End	Size	Genes in Fragment
pS1	73816	77209	3394	(fusA1) hypothetical protein [lp_0077], permease/ drug-metabolite transporter [lp_0078], (GNAT acetyltransferase)
pS2	1004096	1006110	2015	(mae) citC citD (citE)
pS3	2866824	2868276	1453	(pts26BCA)
pS4	3158907	3162902	3996	tal2 [lp_3539], [lp_3540], pts34B, pts34A, (lp_3543)
pG1	3199864	3203643	3780	(clpL), (lox)
pG2	1429291	1430754	1464	Lp_1564 membrane protein, (LyTR family transcriptional regulator)
pG3	1086051	1088519	2469	(rfbC), rfbB, rfbD, (transposase, fragment)
pG4	1024357	1026809	2453	(cydA), cydB, (cydC)



**Fig. 18:** Growth tolerance assay performed in 4% ethanol with a 6 hour exposure for *E. coli* MG1655::rpoD cells bearing the HSP plasmid pZS-C<sup>L</sup>G<sup>L</sup>-ES<sup>N</sup> and the library plasmids shown above. **(A)** % viable cell change measured as the ratio of the number of viable cells after 6 hour exposure to initial viable cell count. **(B)** Cell density after 6 hour exposure. Data reported is the average of 3 biological replicates with error bars indicates standard error.

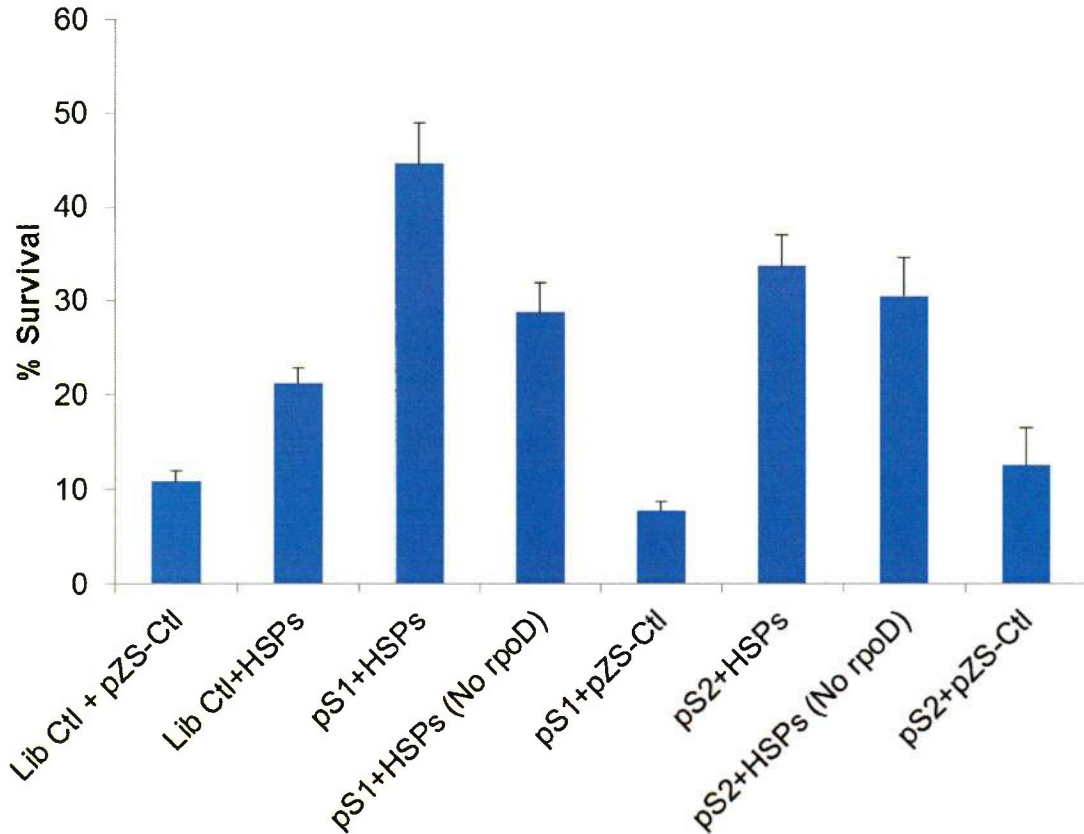


**Fig. 19:** Survival tolerance assay performed in 8% ethanol with a 24 hour exposure for *E. coli* MG1655::rpoD cells bearing the HSP plasmid pZS-C<sup>L</sup>G<sup>L</sup>ES<sup>N</sup> and the library plasmids shown above. Percent survival reported as the ratio of the number of viable cells after 24 hour exposure to initial viable cell count. Data reported is the average of 3 biological replicates with error bars indicates standard error.

The strain expressing the plasmid pG2 [MG1655::rpoD(pZS-C<sup>L</sup>G<sup>L</sup>ES<sup>N</sup>, pG2)] was the only strain tested to produce a statistically significant increase in cell density during the growth screening reaching a final average  $A_{600}$  of 0.822 compared to the control average of 0.684 ( $p < 0.001$ ). However, the strains expressing pG1 and pG2 produced statistically significant increase in viable cell counts during the assay ( $p = 0.031$  and  $p < 0.001$  respectively) with the pG2 expressing strain having the higher percent change of 566% compared to the control strains 310% change. From the strains selected for improved survival only the cells expressing pS1 and pS2 produced significant improvements in percent survival, 44.5% and 33.7% respectively compared to the plasmid controls 21.3% ( $p < 0.001$  for both).

The effect of the identified beneficial plasmids for survival when expressed without the HSP plasmid and without the *L. plantarum* rpoD was also examined (Fig. 20). Without expression of the *L. plantarum* rpoD pS1 produced only a small benefit over the plasmid control when expressed with the HSP plasmid, 28.8% survival versus 44.5%, when expressed with rpoD. The plasmid pS2 produced an equivalent improvement with and without the *L. plantarum* rpoD (33.7% versus 30.5%) indicating that the effect from plasmid pS2 is not dependent upon the *L. plantarum* transcription factor. Expression of either plasmid without the HSP plasmid produced a lower tolerance than expression of the HSP plasmid with a plasmid

control and tolerance comparable to the double plasmid control strain. This indicates that neither plasmid offers significant benefit without the overexpression of the *E. coli* HSPs.



**Fig. 19:** Survival tolerance assay performed in 8% ethanol with a 24 hour exposure for *E. coli* MG1655::rpoD or MG1655 cells bearing the plasmid combinations indicated above. HSPs refers cells bearing the plasmid pZS-C<sup>L</sup>G<sup>L</sup>ES<sup>N</sup>, Lib Ctl is the plasmid control for the library, (No RpoD) indicates that the strain MG1655 was used. Percent survival reported as the ratio of the number of viable cells after 24 hour exposure to initial viable cell count. Data reported is the average of 3 biological replicates with error bars indicating standard error

**VI. Work Plan:** this three-year grant, originally scheduled for completion in September of 2012, was extended without additional costs for 6 months into early 2013 and is now completed. No funds are available for now to continue this line of work. Below are a set of ideas as to what we would like to pursue to extend and enhance the impact of our findings and open new metabolic engineering and synthetic-biology venues in tolerance engineering.

**Related to Objective 1:**

**New directed evolution strategy to build complex *Escherichia coli* phenotype by combining chromosome DNA from *L. plantarum*.** A recent study has shown that the full-length Pac prophage protein RecE and its partner RecT mediate highly efficient linear-linear homologous recombination in *E. coli* (Fu et al., 2012). This discovery provides a possibility that two or more linear DNA fragments flanking with the same homologous region can be assembled to construct heterologous pathways in *E. coli*. In our project, the sheared *L. plantarum* chromosome DNA will be first ligated into appropriate vector, and the vector backbone will be linearized to produce fragments harboring foreign DNA. Subsequently, these fragments and another designed vector backbone flanking with the same homologous region will be assembled in *E. coli* mediated by RecET to create a plasmid library. The library will be screened for ethanol tolerance.

**Characterization of sequenced gene fragment.** The whole identified insert, as well as each part of the fragment and their combinations, will be expressed in *E. coli* to examine their functionalities. The constitutive bla promoter will be employed for gene expression.

**Cohesive ligation.** Continue test cohesive ligation and determine the efficiency of both blunt (previously used) and cohesive ligation.

**Sequencing of the inserted fragment:** currently we are trying to sequence the whole chromosome DNA of the optimized high ethanol tolerance strain EC100gmlp3 using Illumina deep sequencing.

**Related to Objective 2:**

Using unrestricted funds available to the PI, we will continue development of our semi-synthetic stress response system through optimization of the genes identified both through rational approaches and in library selections. Completion of the studies on the identified genes will be completed including analysis of all necessary controls under both growth and survival conditions. Combinations of the identified fragments will also be examined to determine if improvements in survival under ethanol stress can be complemented by genes imparting improved growth to produce a superior platform strain for production. Further exploration of interacting partners may also be done utilizing the library based enrichment approaches described here.

Additionally, we will continue examinations the effect of the overexpression of ClpB, GrpE, and GroESL and the *L. plantarum* genes identified including *murA2* on production of ethanol using the ethanologenic strain *E. coli* KO11 containing has the *Zymomonas mobilis* genes *pdg*, *adhB*, and *cat* integrated into its chromosome using bioreactor based production assays. We will examine fermentations first using our plasmid based approach and we will also integrate the successful gene combinations into the chromosome to reduce the additional

metabolic burden produced by the plasmid based expression system and to develop a more robust platform strain for future engineering.

#### **VII. Major Problems/Issues (if any).**

No major problems.

#### **VIII. Technology Transfer**

No activity this reporting period

#### **IX. Foreign Collaborations and Supported Foreign Nationals**

No foreign collaborations. In YR1, because we could not recruit experienced graduate students immediately when the funding decision for this grant was made in the summer of 2009, the best approach for making progress was to hire a postdoctoral fellow (Dr. Changhao Bi; a citizen of China - foreign national) who had applied for a position in our lab. We then recruited a first year PhD student (Mr. Kyle Zingaro, a US citizen who works on Objective 2). In 2012 a postdoctoral fellow (Dr. Yongbo Yuan), a citizen of China (foreign national) joined our group and has been working on Objective 1 of this project. Furthermore, a PhD student (Dr. Sergios Nicolaou; a citizen of Cyprus (foreign national)) has been involved with this project since its conception almost 4 years ago. Dr. Nicolaou has provided methods guidance and intellectual support to Dr. Bi, Dr. Yuan, and Mr. Zingaro, although he has not been financially supported by this ONR. Dr. Nicolaou has been actively working on this project on a variant approach to Objective 1 as reported here. Dr. Bi left our laboratory in August of 2011.

#### **X. Productivity**

##### **A. Refereed journal articles**

1. Zingaro KA, Papoutsakis ET. 2012. Toward a semisynthetic stress response system to engineer microbial solvent tolerance. *mBio*. 3(5):e00308-12.
2. Nicolaou SA, Gaida SM, & Papoutsakis, ET. 2012. Exploring the combinatorial genomic space in *Escherichia coli* for ethanol tolerance. *Biotech J*. 7:1337-1345.
3. Zingaro KA, Papoutsakis ET. 2013. GroESL overexpression imparts *Escherichia coli* tolerance to i-, n-, and 2-butanol, 1,2,4-butanetriol and ethanol with complex and unpredictable patterns. *Metab. Eng.* 15:196-205.
4. Zingaro KA, Nicolaou SA, Papoutsakis ET. 2013. Of assays and strains: microbial tolerance to toxic chemicals for industrial bioprocessing. *Trends Biotechnol.* Under Review.
5. Nicolaou SA, Fast AG, Nakamaru-Ogiso E, Papoutsakis ET. 2013. Overexpression of FetA (YbbL) and FetBC (YbbM) enhance resistance to oxidative stress in *Escherichia coli*. *To be submitted.*

**B. Non-refereed publications**

None in this reporting period

**C. Books or Chapters**

Papoutsakis ET & Alsaker KV. 2012. Towards a Synthetic Biology of the Stress-Response and the Tolerance Phenotype. Chapter 7 in: "Systems Metabolic engineering" C. Wittmann and S.Y. Lee (eds.). DOI 10.1007/978-94-007-4534-6\_7. Springer Science+Business Media. Dordrecht. Germany.

**D. Technical reports**

None in this reporting period

**E. Workshops and conferences**

Nicolaou SA & Papoutsakis ET. 2012. Utilizing the CoGeL (Co-Existing/Co-Expressing Genomic Libraries) Technology to develop the complex phenotype of high tolerance to oxidative stress. ACS (American Chemical Society) National Meeting, San Diego, CA, March 25-29, 2012. Contributed.

Zingaro KA & Papoutsakis ET. 2012. Toward a Semisynthetic Stress Response System to Engineer Solvent Tolerance. AIChE Annual Meeting, Pittsburgh, PA. Oct 2012.

**F. Patents**

None in this reporting period

**G. Awards/Honors**

Papoutsakis (PI) was:

- i. Elected Fellow of the American Chemical Society (ACS) (Summer of 2011).
- ii. Appointed (summer of 2011) Editor of Biotechnology Advances (Impact Factor; 9.646; Publisher: Elsevier)
- iii. Appointed (2012) to the Scientific Advisory Board (SAB; <http://bioenergycenter.org/besc/about-besc/people.cfm>) of BESC (DOE BioEnergy Science Center (<http://bioenergycenter.org/besc/index.cfm>) based in Oak Ridge National Laboratory.
- iv. Presented James E. Bailey Award for Outstanding Contributions to the Field of Biological Engineering by the AIChE Society for Biological Engineering

**XI. Award Participants** (receiving salary from this ONR award).

No military personnel worked on this project.

Other research personnel:

1. Mr. Kyle Zingaro (PhD student)
2. Dr. Changhao Bi (postdoctoral)
3. Dr. Yongbo Yuan (postdoctoral)
4. Mr. Jason Coffman (undergraduate)

## REFERENCES CITED

- Andrews, S. C., Robinson, A. K., Rodriguez-Quinones, F., 2003. Bacterial iron homeostasis. *FEMS microbiology reviews*. 27, 215-237.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., Mori, H., 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*. 2, 2006 0008.
- Fu, J., Bian, X. Y., Hu, S. B., Wang, H. L., Huang, F., Seibert, P. M., Plaza, A., Xia, L. Q., Muller, R., Stewart, A. F., Zhang, Y. M., 2012. Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nature Biotechnology*. 30, 440-+.
- Lavrrar, J. L., Christoffersen, C. A., McIntosh, M. A., 2002. Fur-DNA interactions at the bidirectional *fepDGC-entS* promoter region in *Escherichia coli*. *Journal of Molecular Biology*. 322, 983-995.
- Ohta, K., Beall, D. S., Mejia, J. P., Shanmugam, K. T., Ingram, L. O., 1991. Genetic-Improvement of *Escherichia-Coli* for Ethanol-Production - Chromosomal Integration of *Zymomonas-Mobilis* Genes Encoding Pyruvate Decarboxylase and Alcohol Dehydrogenase-Ii. *Applied and Environmental Microbiology*. 57, 893-900.
- Touati, D., Jacques, M., Tardat, B., Bouchard, L., Despied, S., 1995. Lethal Oxidative Damage and Mutagenesis Are Generated by Iron in Delta-Fur Mutants of *Escherichia-Coli* - Protective Role of Superoxide-Dismutase. *Journal of Bacteriology*. 177, 2305-2314.

# Engineering Complex Microbial Phenotypes

E. Terry Papoutsakis, Univ. of Delaware

## Objective:

- Develop *E. coli* cells which are more tolerant to alcohols & toxic chemicals
- Develop evolutionary engineering approaches for generating complex phenotypes

## Approach:

- Identify and characterize autologous and heterologous genomic library fragments imparting improved toxic chemical tolerance
- Improve custom stress-response system developed for tolerance to alcohols

## Accomplishments :

- Enhanced custom stress-response system with heterologous genes (*L. plantarum*) identified in library screenings
- Characterized heterologous genes imparting improved ethanol tolerance
- Examined effects of iron transporters on oxidative stress in *E. coli*

## Impact:

- Significant improvements in alcohol and oxidative stress tolerance over published studies
- Improvement of tolerant platform strains for production strain development

