J. Microbiol. Biotechnol. (2017), 27(6), 1171–1179
https://doi.org/10.4014/jmb.1702.02021

jmb

Global Functional Analysis of Butanol-Sensitive *Escherichia coli* and Its Evolved Butanol-Tolerant Strain S

Haeyoung Jeong[1], Seung-Won Lee[2], Sun Hong Kim[1#], Eun-Youn Kim[3], Sinyeon Kim[4], and Sung Ho Yoon[4*]

[1]*Infectious Disease Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon 34141, Republic of Korea*
[2]*SeqGenesis, Daejeon 34111, Republic of Korea*
[3]*School of Basic Sciences, Hanbat National University, Daejeon 34158, Republic of Korea*
[4]*Department of Bioscience and Biotechnology, Konkuk University, Seoul 05029, Republic of Korea*

Received: February 10, 2017
Revised: March 14, 2017
Accepted: March 18, 2017

First published online
March 24, 2017

**Corresponding author*
Phone: +82-2-450-3761;
Fax: +82-2-450-0686;
E-mail: syoon@konkuk.ac.kr

#Present address: D&P Biotech, Inc., Daegu 41404, Republic of Korea

Supplementary data for this paper are available on-line only at http://jmb.or.kr.

pISSN 1017-7825, eISSN 1738-8872

Butanol is a promising alternative to ethanol and is desirable for use in transportation fuels and additives to gasoline and diesel fuels. Microbial production of butanol is challenging primarily because of its toxicity and low titer of production. Herein, we compared the transcriptome and phenome of wild-type *Escherichia coli* and its butanol-tolerant evolved strain to understand the global cellular physiology and metabolism responsible for butanol tolerance. When the ancestral butanol-sensitive *E. coli* was exposed to butanol, gene activities involved in respiratory mechanisms and oxidative stress were highly perturbed. Intriguingly, the evolved butanol-tolerant strain behaved similarly in both the absence and presence of butanol. Among the mutations occurring in the evolved strain, *cis*-regulatory mutations may be the cause of butanol tolerance. This study provides a foundation for the rational design of the metabolic and regulatory pathways for enhanced biofuel production.

Keywords: Butanol, tolerance, evolution, transcriptome, *Escherichia coli*

Introduction

With global warming and the energy crisis, the biological production of carbon-neutral fuels has received great attention worldwide. Butanol is a promising biofuel owing to its higher energy content and lower corrosivity and hygroscopicity [1]. Members of the *Clostridium* genus, which are natural butanol producers, have been metabolically engineered to increase the productivity [2, 3]. Engineering of well-studied hosts such as *Escherichia coli* and *Saccharomyces cerevisiae* can be suitable for improving the yield and level of tolerance during butanol production [4]. With its rapid and reliable growth in high-cell density cultures and excellence in genetic manipulation, *E. coli* is one of the primary choices for the industrial production of biofuels and biorefineries [5]. Metabolically engineered *E. coli* strains have been developed for the heterologous expression of the butanol biosynthesis pathway [6–9]. The major challenge in butanol production using *E. coli* strains is in increasing their tolerance to butanol. Wild-type *E. coli* strains cannot withstand up to 1% (v/v) butanol [4].

Various approaches have been explored to identify gene candidates for enhancing the butanol tolerance in *E. coli*, such as the random mutagenesis of the global transcription factor [10], the expression of artificial transcription factors [11], genomic library enrichment [12], and the overexpression of chaperonin genes [13]. Although these attempts have yielded an increase in tolerance of up to 1.5% butanol, engineering more robust hosts with enhanced butanol production requires omics analyses and an experimental evolution for a better understanding of the butanol stress response and the identification of potential targets for

butanol production [14, 15].

We previously developed and genome-sequenced the enhanced 1-butanol-tolerant *E. coli* [16] that can grow in 1.3% 1-butanol (v/v), which was derived from wild-type *E. coli* by combining the experimental evolution [17] and mutagenesis through recurrent proton irradiation [18]. A genome sequencing analysis of the evolved cells identified 11 coding mutations and three *cis*-regulatory mutations [3]. The coding mutations occurred in genes involved in chaperonin, cytoplasmic membrane biosynthesis, and ethanol metabolism. Among these genes, *fabF* encoding β-ketoacyl-acyl carrier protein synthase II was reported to change the fatty acid composition of the cytoplasmic membrane, and was thought to increase the membrane fluidity at low temperature [19]. To test whether the evolved *fabF* was the origin of the butanol-tolerant phenotype, the *E. coli* strain with the evolved allele was constructed; however, to our disappointment, the tolerance was hardly enhanced [16]. In this study, the transcriptomes and phenomes of wild-type *E. coli* and its butanol-tolerant evolved cells were compared to better understand the global cellular physiology and metabolism involved in the butanol tolerance, and by extension, to pinpoint the genetic basis for the origin of the butanol-tolerant phenotype.

Materials and Methods

Strains and Culture Conditions

Bacterial cell cultures were performed for *E. coli* C strain (ATCC 8739 or KCTC 2571) and its butanol-tolerant evolved strain (PKH5000, which was deposited as KCTC 42934 and whose raw sequencing data are available as SRX1409823 at Sequence Read Archive) [16]. Cells were grown aerobically in 250 ml flasks containing 50 ml of LB medium in a shaking incubator at 37°C and 250 rpm. When the cell density in OD_{600} reached 1, butanol was added to the medium at a final concentration of 0.9% (v/v). Three replicate cultures were carried out. For the transcriptome analysis, the samples were taken just prior to the butanol treatment (referred to as A_0 for the ancestral strain, and E_0 for the evolved strain) and 30 min after the perturbation (A_1 for the ancestral strain, and E_1 for the evolved strain) (Fig. 1).

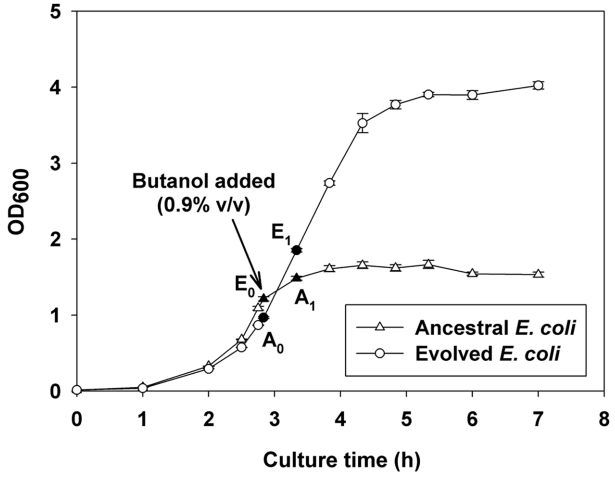

Fig. 1. Growth of the ancestral and evolved *E. coli* strains in LB medium.

The error bar denotes the standard error of the mean from three independent growth curves. Filled symbols are the sampling points for transcriptome analyses.

Transcriptome Analysis

E. coli microarrays were constructed to contain 4,819 70-mer oligonucleotides, which were spotted in duplicates [20]. As the microarray platform GPL7395 used in this study was originally designed for *E. coli* strains of K-12 MG1655 and B REL606, there was sequence similarity between the 70-mer probes and *E. coli* ATCC 8739 from which PKH5000 was derived. Pairwise average nucleotide identities among the three strains as calculated by JSpecies [21] were above 99.97%. All probe sequences were searched against the CDS sequences of ATCC 8739 (NC_010468.1) using the BLAST+ with default parameters [22]. A total of 3,652 probes with maximum 2 bp mismatch for 70 bp probes were chosen for further analysis.

Total RNA extraction and microarray experiments were conducted as previously described [20]. Briefly, the total RNA was isolated using an RNeasy column (Qiagen, USA) according to the manufacturer's instructions. The total RNA from each sample was labeled with Cy3 or Cy5, and two-color array tests were applied to three sets of pairwise comparisons (A_1 *vs.* A_0, E_1 *vs.* E_0, and E_0 *vs.* A_0). The labeled RNA was hybridized to the microarrays. After hybridization and washing, the microarrays were scanned using a GenePix 4000B scanner (Molecular Devices, USA). Array tests were conducted using three biological replicates for each comparison.

The preprocessing and quantile-normalization of the expression intensities were conducted using Bioconductor (ver. 2.12) [23] with the limma package [24]. The log2-transformed transcription ratios of the ancestral strains (A_1/A_0) and evolved strains (E_1/E_0) were calculated. Their false discovery rates (FDRs) were computed from the *p*-values obtained through Benjamin and Hochberg's multiple testing correction. A gene ontology (GO) analysis was conducted using the EMBL STRING database [25]. The array data were deposited in the NCBI Gene Expression Omnibus under the accession number GSE47589.

Analysis of Functional Association Gene Network Combined with Transcriptome Data

The integrated functional association network of genes was inferred using the EMBL STRING database [25]. The minimum score required for the interaction between genes was set as a

default value of 0.4. The nodes of the network were mapped with a transcript change of E_1/A_1 and the FDR of the differences between the relative transcript levels of A_1/A_0 and E_1/E_0. The resulting network was visualized using Cytoscape [26] and explored using the Gaggle framework [27].

Phenome Analysis

Phenotype microarray (PM) tests on the ancestral and evolved strains were carried out as previously described [20]. The PM plates (Biolog Inc., USA) consist of twenty 96-well microplates containing different sources of carbon (PM1 and PM2), nitrogen (PM3), phosphorus and sulfur (PM4), auxotrophic supplements (PM5 through PM8), or salts (PM9). The PM10 plate tests the pH stress, and plates PM11 through PM20 contain inhibitory compounds such as antibiotics, antimetabolites, and other inhibitors. Cells were grown on BUG+B agar overnight at 37°C. Colonies were picked from the agar surface and suspended in an inoculating fluid containing the indicator dye tetrazolium violet. IF-0 inoculating fluid was used for plates PM1 through PM8, and IF-10 fluid was used for plates PM9 through PM20. Disodium succinate and ferric citrate were added to the inoculation solutions of plates PM3 through PM8. All of the PM plates were inoculated with cell suspensions at 100 µl/well and incubated at 37°C for 24 h in an OmniLog incubator (Biolog Inc., USA). The raw kinetic values of the cell growth were imported from OmniLog-PM software, and were then analyzed using the opm R package [28].

Results

Bacterial Cell Growth

To evaluate the range of 1-butanol concentrations inhibitory to the ancestral and evolved *E. coli* strains, the growth of ancestral and evolved strains was monitored in LB media with different concentrations of butanol (added to the medium at the beginning) (Fig. S1). The ancestral and evolved strains grew slightly in the medium with concentrations of butanol at 1.0% (v/v) and 1.3%, respectively. For the transcriptome analysis, the 1-butanol concentration of 0.9% (v/v) was selected to be the highest inhibitory concentration ensuring the cell growth of both strains after butanol treatment.

The ancestral and evolved *E. coli* strains grew in a similar fashion prior to the addition of 0.9% (v/v) butanol to the culture media (Fig. 1). Upon the butanol addition, the evolved cells continued growing exponentially up to 4.0 in OD_{600}. However, the growth of the ancestral strain was rapidly inhibited and reached the stationary phase at 1 h after perturbation. As the growth stage itself results in global dynamic changes in transcriptome structure [20, 29], cells prior to butanol treatment and those after butanol treatment for 30 min were subjected to microarray experiments to minimize the environmental factors other than butanol treatment.

Transcriptomic Differences

All transcript ratios and their FDR are tabulated in Supplemental Table S1. The transcript abundances were compared between the ancestral (denoted as A) and evolved (E) cells cultured before (0) and after (1) butanol treatment (Fig. 2). Among the 25 functional categories from the Clusters of Orthologous Groups (COGs) [30], genes belonging to two categories (energy production and conversion (FDR: 0.042), and translation, ribosomal structure, and biogenesis (FDR: 1.4E-08)) were significantly differently expressed. During butanol treatment, down-regulated genes in both the ancestral and evolved cells (region I in Fig. 2) included *malK-lamB-malM*, which are involved in the transport of maltose, and *rpl* and *rps* operons encoding ribosomal proteins; however, their levels were greatly reduced in the evolved cells compared with those in the ancestral cells. Up-regulated genes in

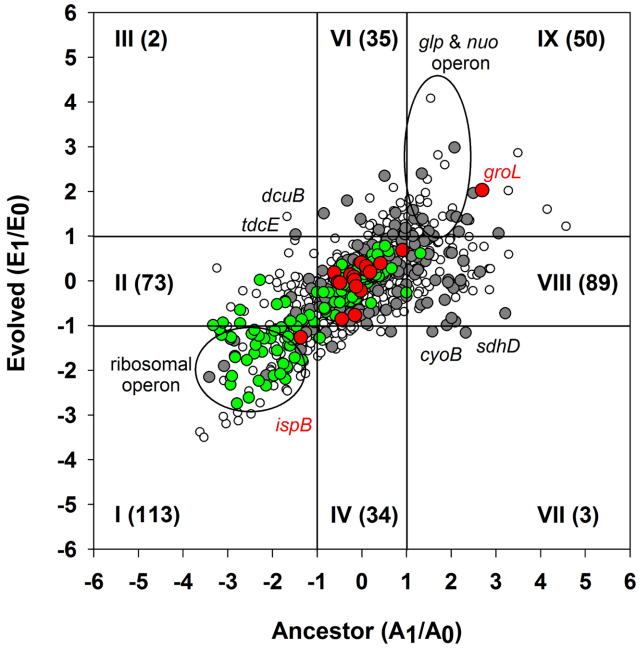

Fig. 2. Scatter plot of the transcript ratios, before and after butanol treatment, of the ancestral (A_1/A_0) and evolved (E_1/E_0) strains.

On both axes of the log2-transformed ratios, the lines are positioned at ±1.0. The numbers of the genes, and the featured genes within each ratio range, are denoted. Each gene was colored according to the COG functional category: energy production and conversion (grey); translation, ribosomal structure, and biogenesis (green); and other functions (white). Mutated genes in the evolved strain are colored red.

both strains (region IX in Fig. 2) were those encoding NADH:ubiquinone oxidoreductase (*nuo* operon), anaerobic glycerol-3-phosphate dehydrogenase (*glpABC*), and chaperonins (*groL*, *groS*, *dnaK*, and *clpB*). Interestingly, during butanol treatment, *dcuB* (encoding the anaerobic C_4-dicarboxylate transporter) and *tdcE* (encoding 2-ketobutyrate formate-lyase involved in the anaerobic degradation of L-threonine into propionate [31]) were up-regulated in the evolved cells and down-regulated in the ancestral cells (region III in Fig. 2), whereas the opposite expression pattern was observed for *cyoB* (encoding cytochrome *o* ubiquinol oxidase subunit) and *sdhCD* (encoding succinate dehydrogenase) (region VII in Fig. 2). Among the mutated genes in the evolved cells, the GroEL chaperonin gene (*groL*) and octaprenyl diphosphate synthase gene (*ispB*), which are essential for the synthesis of the side chain of isoprenoid quinone, were substantially up-regulated and down-regulated in both strains, respectively.

A total of 581 genes showing significant relative transcript changes between the evolved (E_1/E_0) and ancestral (A_1/A_0) cells (FDR < 0.05) were regarded as differentially expressed genes (DEGs) (Table S2). GO enrichment analysis showed that significantly enriched biological processes in DEGs (FDR ≤ 1.65E-09) are associated with aerobic respiration and the oxidation-reduction process (Table S3).

Analysis of the Functional Association Network of Mutated Genes

Previously, we reported 15 point mutations in the evolved cells - ten amino acid substitutions, one frameshift, and four nucleotide substitutions in intergenic regions [16, 18]. Nineteen genes associated with the mutations were tested to infer their integrated functional association network (Fig. S2). The gene network had significantly more interactions among themselves than expected (protein-protein interaction enrichment *p*-value: 0.0286), which indicates that member genes are at least partially biologically connected, as a group. The six-membered group contained genes encoding inner membrane transporters of the DedA family (*yabI* and *yqjA*), transcriptions factors of the AraC family (*araC* and *rob*), the DNA-binding transcriptional dual regulator (*exuR*), and a conserved protein (*creA*). In another group, four genes involved in protein homeostasis (*groL* and *hslV*) and bacterial morphology (*wecA* and *ftsA*) were connected.

Phenomic Differences

PM tests on the ancestral and evolved strains revealed little phenotypic differences between the two strains (Table S4 and Fig. S3). The areas under the curve (AUCs) of the two strains in each of the 1,920 PM wells were calculated using the opm R package [28]. The difference in AUCs for the evolved and ancestral cells was calculated, and its threshold value (± 2,000) supported the cell growth of only one of the two strains. Cell growth was observed for the evolved strains in 5 PM wells, and for the ancestral strains in 18 PM wells (Table S4). Out of 192 different carbon sources tested, only the utilization of bromosuccinic acid was different in the two strains: growth of the ancestral cells and no growth of the evolved cells. Interestingly, the two strains showed different susceptibility to antibiotics targeting biosynthesis of the bacterial cell membranes (cefoxitin, cefazolin, colistin, and piperacillin).

Discussion

Ancestral (Butanol-Sensitive) *E. coli* Compensates for the Energy Loss Caused by the Disruption of the Respiratory Mechanisms and Oxidative Stress

Cells coordinate various cellular stress responses in response to stressful environmental perturbations. Previous analyses of the transcriptome and proteome of wild-type *E. coli* showed that butanol exposure causes complex stress responses, such as a perturbation of the respiratory functions, oxidative stress, acid stress, heat shock, and cell envelope stress [14]. The cellular responses to butanol exposure in the ancestral and evolved strains were quite different in the process of aerobic and anaerobic respiration (Fig. 3). Upon butanol exposure to the ancestral cells, highly expressed genes included those encoding the enzymes for NADH dehydrogenases (*nuo* operon), cytochrome *o* oxidase (*cyo* operon), and the majority of the TCA cycle genes, whereas those involved in the fermentative pathway were down-regulated. The up-regulation of genes involved in aerobic respiration (*nuo*, *cyo*, and *sdh* operons and TCA cycle genes) was also observed from previous omics analyses of wild-type *E. coli* exposed to butanol [14], which can be attributed to an increased energy demand compensating for the energy loss caused by the disruption of respiratory mechanisms and oxidative stress [14, 32]. The increased aerobic respiration seemed to result in increased oxidative stress, which can be evidenced by the high expression of *fumC* encoding one of the fumarase isozymes during the TCA cycle. FumC is known to be insensitive to oxidative damage, and is highly expressed under oxidative stress [33]. Interestingly, two gene sets, *glpABC* and *glpD*, encoding *sn*-glycerol-3-phosphate dehydrogenase under aerobic and anaerobic conditions, respectively, were highly up-regulated in both strains. The

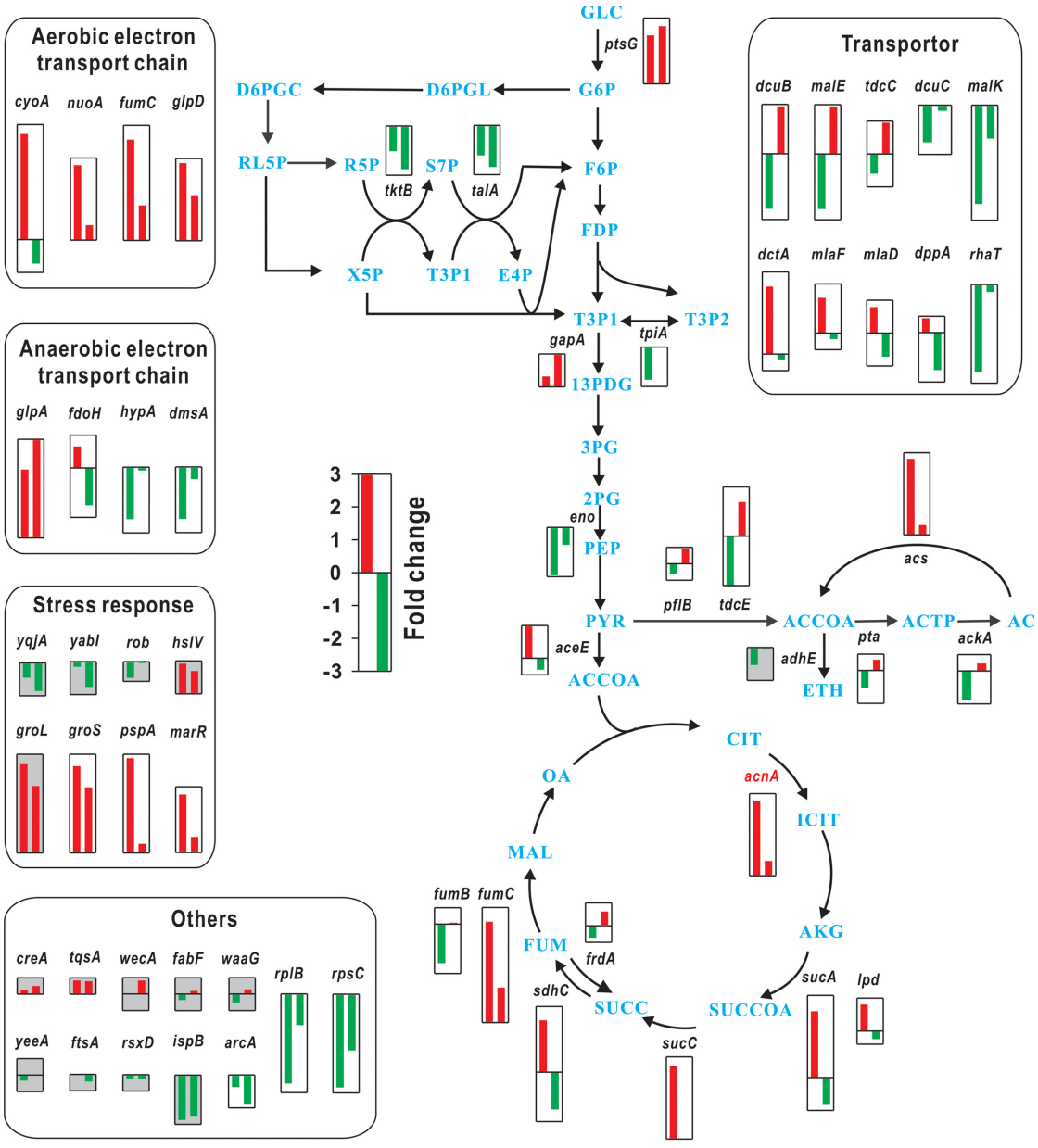

Fig. 3. Comparison of transcript changes under the presence or absence of butanol in the ancestral and evolved *E. coli* strains.
In each of the thumbnail graphs, the left and right bars in the x-axis are ancestral strain and evolved strain, respectively. The y-axes show the log2-transformed transcript ratio (A_1/A_0 or E_1/E_0) of each of the differentially expressed genes (white background) and mutated genes (grey background). For operons, the first operon genes are shown. Metabolites: (13PDG) 1,3-*bis*-phosphoglycerate; (2PG) 2-phosphoglycerate; (3PG) 3-phosphoglycerate; (AC) acetate; (ACCOA) acetyl-CoA; (ACTP) acetyl-phosphate; (AKG) α-ketoglutarate; (CIT) citrate; (D6PGC) D-6-phosphate-gluconate; (D6PGL) D-6-phosphate-glucono-delta-lactone; (E4P) erythrose 4-phosphate; (ETH) ethanol; (F6P) fructose 6-phosphate; (FDP) fructose 1,6-diphosphate; (FUM) fumarate; (G6P) glucose 6-phosphate; (GLC) glucose; (ICIT) isocitrate; (MAL) malate; (OA) oxaloacetate; (PEP) phosphoenolpyruvate; (PYR) pyruvate; (R5P) ribose 5-phosphate; (RL5P) ribulose 5-phosphate; (S7P) sedo-heptulose; (SUCC) succinate; (SUCCOA) succinate-CoA; (T3P1) glyceraldehyde 3-phosphate; (T3P2) dihydroxyacetone phosphate; (X5P) xylulose-5-phosphate.

respiratory enzymes are associated with proton motive force (PMF) generation by oxidizing glycerol-3-phosphate into dihydroxyacetone phosphate. The compromised elements of glycerol-3-phosphate in a phospholipid membrane are probably used to compensate for the PMF loss. In line with this, the down-regulation of *rhaT* (encoding L-rhamnose/

proton symporter) can be explained through the necessity of saving the PMF.

Cell membranes are the primary stress-defense element, and their integrity is essential for bacterial survival. Because butanol can disrupt the integrity of the cell membranes and dissipate the PMF [34], cells need to cope with the envelope stress. Among the major envelope stress response pathways (Bae, Cpx, Psp, and Rcs) [35], *pspABCD* for a phage shock protein (Psp) response was activated. The primary role of the Psp response is to maintain the PMF under membrane stress conditions that perturb the integrity of the inner membrane [36]. The *groESL* chaperonin genes were substantially activated upon butanol exposure, the over-expression of which was reported to reduce the butanol toxicity and improve butanol production in *C. acetobutylicum* [13]. The *marR* of the first gene of the *marRAB* operon, involved in multiple antibiotic resistance, was up-regulated. It should be noted that our observation that butanol treatment caused disruption in the respiratory mechanisms and oxidative stress is in good agreement with the previous report using M9 medium [14], which implies the transcriptomic changes in this study was a genuine response caused by the butanol treatment.

Evolved (Butanol-Tolerant) *E. coli* Behaves Similarly in the Absence and Presence of Butanol

In the evolved cells exposed to butanol, the expression of those genes involved in aerobic respiration and the fermentative pathway did not change appreciably (Fig. 3). Some transporters showed interesting gene expression patterns. Two adjacent and divergently transcribed operons, *malEFG* and *malK-lamB-malM*, involved in the transport of maltose and maltodextrins into the cell, were differently expressed during butanol treatment of the evolved cells: down-regulation of *malK-lamB-malM* and up-regulation of *malEFG*. DctA and DcuB function as co-sensors with the sensor kinase DcuS for C_4-dicarboxylate uptake under aerobic and anaerobic conditions, respectively [37]. The encoding genes, *dctA* and *dcuB*, were up-regulated during butanol treatment in the ancestral and evolved cells, respectively (Fig. 3).

As butanol stress invokes responses common to other stress responses [14], it can be postulated that the evolved butanol-tolerant *E. coli* is likely to show a better degree of fitness under various stressful conditions over the ancestral butanol-sensitive *E. coli*. However, the two strains did not show noticeable phenotypic differences in a variety of stressful conditions caused by osmolarity, pH, or exposure to inhibitory compounds. This means the acquired tolerance was refined to overcome butanol (or solvent) stress, and was not expanded to acquire a more general tolerance mechanism.

The *cis*-Regulatory Mutations Might be the Cause for the Butanol Tolerance in the Evolved Strain

Genomic reorganization is coupled to the evolution of new transcriptional elements [29]. Among the mutated genes, six genes (*adhE*, *groL*, *waaG*, *yabI*, *yfiF*, and *yqjA*) were identified as DEGs (Table 1). The expression patterns of genes associated with mutations were similar in both strains exposed to butanol treatment. Whereas the relative transcript changes before and after butanol treatment between the two strains were moderately correlated for DEGs ($r = 0.56$) (Fig. 2), they were highly correlated for 19 genes associated with the mutations ($r = 0.88$, bootstrapped p-value = 0.0086). This might imply that gene regulatory mechanisms for the mutated genes were not much affected by the point mutations.

To investigate whether the mutations in the evolved strain affected their expression levels during butanol treatment, we compared their transcript levels between the two strains by considering the strain-specific expression pattern (E_0/A_0 and E_1/A_1). During the butanol treatment, compared with the ancestral strains, the high transcript levels in the evolved cells were maintained for *yqjA* and *yabI* (encoding membrane proteins belonging to the DedA protein family), *adhE* (encoding ethanol oxidoreductase during anaerobic growth), *waaG* (formerly *rfaG*; encoding lipopolysaccharide core glycosyltransferase), and *yfiF* (encoding predicted methyltransferase). YqjA and YabI are known to have roles in maintaining the PMF across the cytoplasmic membrane in *E. coli* [38]. In particular, YqjA is required for general envelope maintenance and stabilization of the PMF [39], and its deletion causes an activation of the major envelope stress response pathways, Cpx, Psp, Bae, and Rcs [40]. The high transcript levels in the ancestral cells were observed in a GroEL chaperonin gene (*groL*) and *rob*, encoding a global regulator functioning in resistance to antibiotics, organic solvents, and superoxides [41]. The deletion of the *rob* gene increases the susceptibility to organic solvents, whereas an overexpression of *rob* increases the tolerance to organic solvents as well as the resistance to a variety of antibiotics and superoxide-generating compounds [42].

The strain-specific expression of genes associated with *cis*-regulatory mutations is quite intriguing. Although changes in the *cis*-regulatory elements are an important genetic basis for phenotypic diversity, their causal mutations

Table 1. Strain-specific transcript changes of genes associated with point mutations.

Clone[a]		Genome position	Mutation type[b]	Gene[c]	FDR[d]	Log_2 (mRNA ratio)	
PKH21	PKH5000					E_0/A_0	E_1/A_1
■	■	643,043	Intergenic [*yqjA*(←)/*exuR*(←)]	*yqjA*[e]	0.014	1.67	1.27
				exuR	0.742	0.35	0.57
■	■	1,195,440	Intergenic [*yfiF*(→)/*ung*(←)]	*yfiF*	0.036	-0.15	0.65
				ung	0.132	-0.27	0.02
■	■	3,928,267	Intergenic [*yabI*(←)/*araC*(←)]	*yabI*[e]	0.048	0.81	0.20
				araC	0.937	0.01	0.03
■	■	4,008,590	Intergenic [*creA*(←)/*rob*(→)]	*rob*[e]	0.138	-2.73	-2.28
				creA[e]	0.140	0.01	0.14
	■	84,784	Missense	*waaG*	0.018	-0.05	0.34
	■	554,309	Missense	*ispB*	0.598	0.08	0.18
	■	1,815,193	Missense	*yeeA*	0.245	-0.04	0.13
	■	2,207,491	Missense	*rsxD*	0.965	-0.07	-0.07
	■	2,242,081	Nonsense	*tqsA*	0.967	0.20	0.18
■	■	2,627,410	Missense	*adhE*	0.040	0.22	0.72
■	■	2,757,748	Missense	*fabF*	0.190	0.10	0.37
	■	3,894,759	Missense	*ftsA*	0.246	-0.08	-0.27
	■	4,225,623	Missense	*groL*	0.031	-0.30	-0.96
	■	4,513,154	Missense	*hslV*	0.072	-0.12	-0.34
■	■	4,661,985	Frameshift	*wecA*	0.514	-0.04	0.38

[a]Grey shading indicates mutations that occurred in evolved *E. coli* of PKH21 [18] and PKH5000 [16].

[b]Mutations in intergenic regions have two flanking genes, with their direction of transcription shown in the square brackets.

[c]Genes associated with the corresponding mutation.

[d]False discovery rate of a gene, calculated based on A_1/A_0 and E_1/E_0.

[e]Possible *cis*-regulatory mutations.

for specific phenotypic changes are difficult to identify [43, 44]. In this study, genes associated with *cis*-regulatory mutations (*yqjA*, *yabI*, and *rob*) showed substantial strain-specific transcript levels (Table 1 and Fig. S2). Interestingly, they were mutations that occurred during the early phase of cyclic selection (PKH21 strain) [18], which agrees with a previous study in which the fitness was considerably enhanced through mutations occurring over the early generations during the experimental evolution with *E. coli* [17]. All of these observations might imply that butanol tolerance in the evolved cells could be acquired by modified expression levels of genes associated with the *cis*-regulatory mutations.

In this study, we analyzed the transcriptome and phenome of the ancestral and evolved *E. coli* strains to better understand the tolerance mechanism and identify causal mutations for butanol tolerance. The omics analyses and the identified gene targets should be valuable in understanding the mechanism of butanol tolerance and developing microbial cell factories for the mass production of the biofuel. Moreover, this study would be critical to interpreting omics analysis of an industrial butanol-producing setup, such as the anaerobic or microaerobic conditions.

Acknowledgments

The work of SHY was supported by the National Research Foundation of Korea through the Technology Development Program to Solve Climate Changes on Systems Metabolic Engineering for Biorefineries (2012M1A2A2026559), and the Korean Ministry of Agriculture, Food, and Rural Affairs through the Strategic Initiative for Microbiomes in Agriculture and Food (916006-2). The work of HJ was supported by the Korean Ministry of Science, ICT, and

Future Planning through the Proton Engineering Frontier Project and the Microbial Genomics and Applications Center Program, and the KRIBB Research Initiative Program.

References

1. Fortman JL, Chhabra S, Mukhopadhyay A, Chou H, Lee TS, Steen E, *et al*. 2008. Biofuel alternatives to ethanol: pumping the microbial well. *Trends Biotechnol.* **26:** 375-381.
2. Lutke-Eversloh T, Bahl H. 2011. Metabolic engineering of *Clostridium acetobutylicum*: recent advances to improve butanol production. *Curr. Opin. Biotechnol.* **22:** 634-647.
3. Jang YS, Lee JY, Lee J, Park JH, Im JA, Eom MH, *et al.* 2012. Enhanced butanol production obtained by reinforcing the direct butanol-forming route in *Clostridium acetobutylicum*. *MBio* **3:** e00314-12.
4. Knoshaug EP, Zhang M. 2009. Butanol tolerance in a selection of microorganisms. *Appl. Biochem. Biotechnol.* **153:** 13-20.
5. Yoon SH, Jeong H, Kwon S-K, Kim JF. 2009. Genomics, biological features, and biotechnological applications of *Escherichia coli* B: "Is B for better?!", pp. 1-17. *In* Lee SY (ed.). *Systems Biology and Biotechnology of Escherichia coli*. Springer, Berlin, Germany.
6. Atsumi S, Cann AF, Connor MR, Shen CR, Smith KM, Brynildsen MP, *et al*. 2008. Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab. Eng.* **10:** 305-311.
7. Gulevich AY, Skorokhodova AY, Sukhozhenko AV, Shakulov RS, Debabov VG. 2012. Metabolic engineering of *Escherichia coli* for 1-butanol biosynthesis through the inverted aerobic fatty acid beta-oxidation pathway. *Biotechnol. Lett.* **34:** 463-469.
8. Shen CR, Liao JC. 2008. Metabolic engineering of *Escherichia coli* for 1-butanol and 1-propanol production via the keto-acid pathways. *Metab. Eng.* **10:** 312-320.
9. Dong H, Zhao C, Zhang T, Lin Z, Li Y, Zhang Y. 2016. Engineering *Escherichia coli* cell factories for *n*-butanol production. *Adv. Biochem. Eng. Biotechnol.* **155:** 141-163.
10. Zhang H, Chong H, Ching CB, Song H, Jiang R. 2012. Engineering global transcription factor cyclic AMP receptor protein of *Escherichia coli* for improved 1-butanol tolerance. *Appl. Microbiol. Biotechnol.* **94:** 1107-1117.
11. Lee JY, Yang KS, Jang SA, Sung BH, Kim SC. 2011. Engineering butanol-tolerance in *Escherichia coli* with artificial transcription factor libraries. *Biotechnol. Bioeng.* **108:** 742-749.
12. Reyes LH, Almario MP, Kao KC. 2011. Genomic library screens for genes involved in *n*-butanol tolerance in *Escherichia coli*. *PLoS One* **6:** e17678.
13. Tomas CA, Welker NE, Papoutsakis ET. 2003. Overexpression of *groESL* in *Clostridium acetobutylicum* results in increased solvent production and tolerance, prolonged metabolism, and changes in the cell's transcriptional program. *Appl. Environ. Microbiol.* **69:** 4951-4965.
14. Rutherford BJ, Dahl RH, Price RE, Szmidt HL, Benke PI, Mukhopadhyay A, *et al*. 2010. Functional genomic study of exogenous *n*-butanol stress in *Escherichia coli*. *Appl. Environ. Microbiol.* **76:** 1935-1945.
15. Atsumi S, Wu TY, Machado IM, Huang WC, Chen PY, Pellegrini M, *et al*. 2010. Evolution, genomic analysis, and reconstruction of isobutanol tolerance in *Escherichia coli*. *Mol. Syst. Biol.* **6:** 449.
16. Jeong H, Kim S, Han S, Kim M, Lee K. 2012. Changes in membrane fatty acid composition through proton-induced *fabF* mutation enhancing 1-butanol tolerance in *E. coli*. *J. Korean Phys. Soc.* **61:** 227-233.
17. Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, *et al.* 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461:** 1243-1247.
18. Jeong H, Han J. 2010. Enhancing the 1-butanol tolerance in *Escherichia coli* through repetitive proton beam irradiation. *J. Korean Phys. Soc.* **56:** 2041-2045.
19. Garwin JL, Klages AL, Cronan JE Jr. 1980. Beta-ketoacyl-acyl carrier protein synthase II of *Escherichia coli*. Evidence for function in the thermal regulation of fatty acid synthesis. *J. Biol. Chem.* **255:** 3263-3265.
20. Yoon SH, Han MJ, Jeong H, Lee CH, Xia XX, Lee DH, *et al.* 2012. Comparative multi-omics systems analysis of *Escherichia coli* strains B and K-12. *Genome Biol.* **13:** R37.
21. Richter M, Rossello-Mora R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. USA* **106:** 19126-19131.
22. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, *et al*. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10:** 421.
23. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, *et al.* 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5:** R80.
24. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43:** e47.
25. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, *et al.* 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43:** D447-D452.
26. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, *et al.* 2007. Integration of biological networks and gene expression data using Cytoscape. *Nat. Protoc.* **2:** 2366-2382.
27. Shannon PT, Reiss DJ, Bonneau R, Baliga NS. 2006. The Gaggle: an open-source software system for integrating bioinformatics software and data sources. *BMC Bioinformatics* **7:** 176.

28. Vaas LA, Sikorski J, Hofner B, Fiebig A, Buddruhs N, Klenk HP, *et al*. 2013. opm: an R package for analysing OmniLog® phenotype microarray data. *Bioinformatics* **29:** 1823-1824.
29. Yoon SH, Reiss DJ, Bare JC, Tenenbaum D, Pan M, Slagel J, *et al*. 2011. Parallel evolution of transcriptome architecture during genome reorganization. *Genome Res.* **21:** 1892-1904.
30. Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28:** 33-36.
31. Hesslinger C, Fairhurst SA, Sawers G. 1998. Novel keto acid formate-lyase and propionate kinase enzymes are components of an anaerobic pathway in *Escherichia coli* that degrades L-threonine to propionate. *Mol. Microbiol.* **27:** 477-492.
32. Volkers RJ, Ballerstedt H, Ruijssenaars H, de Bont JA, de Winde JH, Wery J. 2009. *TrgI*, toluene repressed gene I, a novel gene involved in toluene-tolerance in *Pseudomonas putida* S12. *Extremophiles* **13:** 283-297.
33. Park SJ, Gunsalus RP. 1995. Oxygen, iron, carbon, and superoxide control of the fumarase *fumA* and *fumC* genes of *Escherichia coli*: role of the *arcA*, *fnr*, and *soxR* gene products. *J. Bacteriol.* **177:** 6255-6262.
34. Dunlop MJ. 2011. Engineering microbes for tolerance to next-generation biofuels. *Biotechnol. Biofuels* **4:** 32.
35. Rowley G, Spector M, Kormanec J, Roberts M. 2006. Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. *Nat. Rev. Microbiol.* **4:** 383-394.
36. Darwin AJ. 2005. The phage-shock-protein response. *Mol. Microbiol.* **57:** 621-628.
37. Janausch IG, Zientz E, Tran QH, Kroger A, Unden G. 2002. C_4-dicarboxylate carriers and sensors in bacteria. *Biochim. Biophys. Acta* **1553:** 39-56.
38. Doerrler WT, Sikdar R, Kumar S, Boughner LA. 2013. New functions for the ancient DedA membrane protein family. *J. Bacteriol.* **195:** 3-11.
39. Kumar S, Doerrler WT. 2014. Members of the conserved DedA family are likely membrane transporters and are required for drug resistance in *Escherichia coli*. *Antimicrob. Agents Chemother.* **58:** 923-930.
40. Sikdar R, Simmons AR, Doerrler WT. 2013. Multiple envelope stress response pathways are activated in an *Escherichia coli* strain with mutations in two members of the DedA membrane protein family. *J. Bacteriol.* **195:** 12-24.
41. Jair KW, Yu X, Skarstad K, Thony B, Fujita N, Ishihama A, Wolf RE Jr. 1996. Transcriptional activation of promoters of the superoxide and multiple antibiotic resistance regulons by Rob, a binding protein of the *Escherichia coli* origin of chromosomal replication. *J. Bacteriol.* **178:** 2507-2513.
42. Bennik MH, Pomposiello PJ, Thorne DF, Demple B. 2000. Defining a *rob* regulon in *Escherichia coli* by using transposon mutagenesis. *J. Bacteriol.* **182:** 3794-3801.
43. Wray GA. 2007. The evolutionary significance of *cis*-regulatory mutations. *Nat. Rev. Genet.* **8:** 206-216.
44. Stern DL, Orgogozo V. 2008. The loci of evolution: how predictable is genetic evolution? *Evolution* **62:** 2155-2177.