**Tailoring of global transcription sigma D factor by random
mutagenesis to improve *Escherichia coli* tolerance towards low-pHs**

Xi Gao ¹, Ling Jiang ^{2,*}, Liying Zhu ³, Qing Xu ¹, Xian Xu ¹, He Huang ^{1,*}

¹College of Biotechnology and Pharmaceutical Engineering, Nanjing Tech University,
Nanjing 210009, PR China; ²College of Food Science and Light Industry, Nanjing Tech
University, Nanjing 210009, PR China; ³College of Sciences, Nanjing Tech University,
Nanjing 210009, PR China

Xi Gao: melindag@yeah.net; Ling Jiang: jiangling@njtech.edu.cn;

Liying Zhu: zlyhappy@njtech.edu.cn; Qing Xu: xuqing357@njtech.edu.cn;

Xian Xu: xuxian@njtech.edu.cn; He Huang: biotech@njtech.edu.cn

*Corresponding authors.

Ling Jiang, Email: jiangling@njtech.edu.cn, Tel: +86-25-58139942,
College of Food Science and Light Industry, Nanjing Tech University, Nanjing 210009,
PR China.

He Huang, Email: biotech@njtech.edu.cn, Tel: +86-25-58139942,
College of Pharmaceutical Science, Nanjing Tech University, Nanjing 210009, PR
China.

Abstract

Bioconversion processes of organic acid or acid hydrolysis of raw material for microbial metabolism often suffer limitations as a result of microbial sensitivity in low-pH conditions. We adopted a three-step method called RAndom Insertional-deletional Strand Exchange mutagenesis (RAISE) to engineer the components of global regulator Sigma D factor (RpoD) of *Escherichia coli* to improve its acid tolerance. The best strain Mutant VII was identified from random mutagenesis libraries based on the growth performance, which exhibited much higher growth rate than the control (0.22 h^{-1} vs. 0.15 h^{-1}) at pH as low as 3.17. Combined transcriptome and phenome analysis of *E. coli* was carried out to better understand the global effects of RpoD on the regulatory networks. Our analysis showed that 95 (2.1%) of all *E. coli* genes were induced and 178 (4.0%) genes were repressed, including those for trehalose biosynthesis, nucleotides biosynthesis, carbon metabolism, amino acid utilization, except for acid resistance. Also regulated were the master regulators (ArcA, EvgA, H-NS and RpoS) and gene/operon-specific transcription factors (GadX, GadW, AppY, YdeO, KdgR). These results demonstrated that RpoD acts as global regulator in the growth phase of *E. coli* and consequently improves acid tolerances.

Keywords: *Escherichia coli*; acid tolerance; RpoD; RNA-seq; Phenotype microarray

1. Introduction

Biorefining promises the development of efficient biocatalysis processes for the conversion of renewable biomass resources into building block chemicals that are amenable to further conversions to higher value products as well as large volume commodity chemicals. Nevertheless, many challenges still remain for the economical bio-production of these chemicals. For example, product toxicity issue associated with the bioconversion processes of organic acid is one of the primary challenges based on the use of platform organisms, such as *Escherichia coli*. *E. coli* has been chosen as a model organism for the production of high-value organic acids like succinic acid (Chan et al., 2012), 3-hydroxypropionic acid (Kwak et al., 2013), lactic acid (Mazumdar et al., 2013), and so on, due to its clear genetic background, convenient to be genetically modified as well as good growth property with low nutrient requirements. However, high titers of organic acid in the free acid form during fermentation inhibits both cell growth and acid production of *E. coli* (Chun et al., 2014). While the control of fermentation broth pH by addition of large volumes of alkaline substances (e.g., CaCO_3 , NaOH and NH_4OH) could reduce the negative effects of free acid accumulation but strongly foul downstream processing (Abdel-Rahman et al., 2013). Meanwhile, acid-hydrolysis-mediated pretreatment of raw material to release its component sugars which are suitable for microbial metabolism, in turn impair viability of microbes because of the low pH circumstance (Huang et al., 2009). One model for addressing such challenges involves the improvement of cellular phenotype under low pH or acid

1 stress conditions.

2 Measures have previously been taken to elevate *E. coli* acid tolerance through
3 classical strain engineering approaches such as UV/chemical mutagen treatment,
4 adaptive evolution, which are often time and labour-intensive (Dragosits and
5 Mattanovich, 2013). Metabolic engineering tools, including deletion of *glpK*, *yqhD*,
6 *gldA*, *pta*, *ldhA*, and individual over-expression of *cacKB*, *cscA*, *dhaB*, *dhaR*, *aldH* have
7 also been employed, which are, however, difficult to reach a global phenotype optimum
8 due to the complexity of metabolic landscapes (Chan et al., 2012; Kwak et al., 2013).
9 More recently, transcriptional engineering starts to emerge as a alternative strategy for
10 the improvement of strain tolerance and production phenotypes through multiple,
11 simultaneous perturbations regulated by reprogramming the transcriptome (Alper, 2006).
12 A successful case is reported by Jiang et al. that the global regulator cAMP receptor
13 protein (CRP) has been engineered by error-prone PCR to improve the tolerance of *E.*
14 *coli* towards low-pH and acetate stresses (Basak et al., 2014). The mutant strain almost
15 doubled (0.113 h^{-1}) the growth rate of the control (0.062 h^{-1}) at pH 4.24. However, the
16 growth rate needs to be further improved to better meet the requirements for
17 industrialized production. It has been pointed out that the pattern of genome
18 transcription in bacteria is mainly determined by controlling the utilization of RNA
19 polymerase, the promoter specificity of which was further controlled by several species
20 of RNA polymerase sigma factor (Ishihama, 2010). Earlier work revealed that the
21 global regulator CRP in *E. coli* played an important regulatory role in synthesis of the

stationary-phase specific sigma S factor (RpoS), which was actually responsible for many different stress conditions (Hengge-Aronis, 2002). Most analyses have, however, paid full attention to the strains entering into stationary phase accompanied with a reduction or cessation of growth, and ignored the influence of exponential-phase global regulators on the evolution of host regulatory networks. The sigma D factor, encoded by the *rpoD* gene, plays a primary and major role in transcription of most of the genes that is expressed in exponentially growing phase of *E. coli*, and can partially replace the sigma factor RpoS under many stress conditions (Alper and Stephanopoulos, 2007). In particular, it is reported that *E. coli* is surprisingly acid sensitive in exponential phase when cultured planktonically (Richard and Foster, 2004).

In this work, the *rpoD* gene was subjected to random mutagenesis by Random Insertional-deletional Strand Exchange mutagenesis (RAISE) method, which is based on gene shuffling except the tailing by terminal deoxynucleotidyl transferase (TdT) before the self-priming PCR step (Fujii et al., 2006). The RAISE product was inserted into a low-copy expression vector pKSC by overlap extension PCR (Basak et al., 2014). By taking the approach of mutagenesis and phenotype selection on the *rpoD* gene, the major goal is to elicit a desired cellular phenotype with elevated low pH tolerance. Furthermore, transcriptome and phenome properties as well as quantitative real-time PCR (qRT-PCR) measurement were carried out on the best mutant to investigate the effects of the global regulator RpoD on the changes in gene expression in the heterologous *E. coli* host.

2. Materials and methords

2.1. Strains and media

E. coli DH5 α was applied to routine transformations, as well as RNase-free DNase I, TdT, and restriction enzymes were purchased from TaKaRa (Shanghai, China). Strains were cultured in improved Luria–Bertani (LB) broth at 37°C with 200 rpm orbital shaking or on improved LB agar at 37°C. Earlier studies reported that acid sensitivity of cells is enhanced by incorporation of more than 150 mM sodium chloride (Rowbury et al., 1994), a therefore, 5 g/l NaCl (86 mM) was utilized in LB here. Culture was supplemented with 30 μ g/ml of kanamycin for low-copy-number (~5 copies/cell) host plasmid breeding (pKSC plasmid). The DH5 α strain containing plasmid pKSC-*rpoD* is denoted as a control strain in this study. All remaining chemicals were purchased from Sigma-Aldrich and primers were from Invitrogen.

2.2. Library construction

A low-copy-number plasmid (pKSC) was structured by Rongrong Jiang according to a previously reported protocol with improvement (Zhang et al., 2012). The PCR primers used in this work are displayed in Table 1. Sigma D factor was inspected for a target of RAISE method (Fujii et al., 2006). The *rpoD* operon was amplified from *E. coli* genomic DNA with primers *rpoD*-sense-*Eco*RI and *rpoD*-anti-*Hind*III. The purified PCR products (2,075 bp) were digested with RNase-free DNase I in the buffer

1 containing Tris-HCl (pH 7.0) and MnCl₂. To avoid introduction of too many point
2 mutations in the ensuing self-priming PCR step, MnCl₂ was utilized as a DNase I
3 co-factor to control the size of the digestion products. Determined by agarose gel
4 electrophoresis, the fragments of 100–300 bp were terminated by adding Ethylene
5 Diamine Tetraacetic Acid (EDTA). The addition of TdT was randomly attached to the 3'
6 terminus of the fragments in the solution containing buffer and dNTPs. With a DNA
7 polymerase possessing no proofreading activity, the 3'-tailed purified fragments were
8 reassembled by self-priming PCR.

9 The full-length of the mutated *rpoD* gene (RAISE product) was amplified by PCR,
10 utilising the M13 universal primers. The purified RAISE products were digested by
11 *EcoRI* and *HindIII* overnight, ligated into a digested pKSC, and transformed into *E. coli*
12 DH5 α . The transformed cells were plated onto LB agar containing 30 μ g/ml of
13 kanamycin and scraped off white colonies to create a liquid library. The same procedure
14 (and library size) was used for the three rounds of mutagenesis used in the low pH
15 experiment.

17 2.3. Phenotype selection

18 Phenotype selection and evaluation of samples from the liquid library were chosen
19 based on acid shock method. The random mutagenesis library was inoculated into fresh
20 50-ml LB broth at 37°C with 200 rpm orbital shaking. When strains grew up to the
21 beginning of the exponential phase (OD₆₀₀~0.3), 36% (w/w) HCl was added to the

medium to regulate its pH decreasing from 4.0 to 3.0 for 4 generation and the strains were cultured at 37°C with 200 rpm orbital shaking for 2 h. The strains were then washed with fresh LB broth thrice and spread onto LB-kanamycin agar plates after incubation. After culturing at 37°C, 20 colonies were picked up randomly and their plasmids were isolated by miniprep. DNA sequencing was used to identify the mutations in *rpoD*. To remove any possible interference, the mutant *rpoD* was backcloned into fresh pKSC plasmid, and then transformed back to fresh *E. coli* DH5 α . The strain showing the best improvement of acid tolerance was chosen as the greatest mutant from Round 1 and was respectively cultured under pH decreasing from 7.0 to 3.0. The above procedure was repeated for the three rounds of mutagenesis used in selecting a mutant strain of highest enhancement of acid resistance.

After above experiments, each overnight colony (both mutant and control) was inoculated into fresh improved LB-kanamycin broth to grow up to OD₆₀₀~0.3, and then pH of the mediums were regulated from 3.17, 3.62, 4.15 to 6.86. Strain growth were recorded by monitoring its absorbance at 600 nm.

2.4. Transcriptome analysis by RNA-seq

Overnight strains (both the control and Mutant VII) were inoculated into improved LB-kanamycin broth to grow up to OD₆₀₀~0.3. pH of the mediums were regulated to 4.15 for 2 h and then the strains were incubated into fresh LB-kanamycin broth (pH ~7.0). When they were grown to an OD of 0.6–0.8, their total RNA were extracted with

1 RNeasy Mini Kit (Qiagen, Germany). RNA-Seq services were provided by BGI tech
2 utilising Illumina HiSeq2000 (Marioni et al., 2008). Gene expression difference analysis
3 was mainly represented by the analysis of Gene Ontology and KEGG Pathway. Cluster
4 software and Java Treeview software were utilized to perform cluster analysis of gene
5 expression patterns. The P-values were corrected by FDR(False Discovery Rate) control.
6 To permit statistical confidence in differential gene expression, each sample had three
7 biological replicates.

8 Quantitative real-time PCR (qRT-PCR) was carried out to verify RNA-Seq results
9 by utilising an ABI STEPONE PLUS Real Time PCR System (Basak et al., 2014).
10 Primer Premier 6.0 (PREMIER Biosoft International, Palo Alto, CA, USA) was utilised
11 to design primers of the target genes (Table 1). cDNA synthesis kit (New England
12 Biolabs) was used to reverse transcribe RNA into first-strand cDNA. The PCR
13 amplification was detected by utilising the SYBR green fluorescence dye. 16s rRNA
14 synthesis gene rrsA was utilised as calibrator.

16 2.5. Phenotype microarray analysis

17 Colonies from LB-kanamycin agar of the overnight test strains (both Mutant VII
18 and the control strain) were picked up and inoculated into microplate at a density
19 corresponding to 85% transmittance by using the Biolog turbidimeter (Biondi et al.,
20 2009). PM1 and PM10 consists of preconfigured well arrays with different carbon
21 resources and pH values and an automated instrument that continuously monitors and

records the response of the cells by measuring their growth curves. The microplate (100 μ l/well) were placed in the OmniLog instrument at 37 °C for 48 h. The OD₅₉₅ was measured automatically every 15 min for 48 h using an OmniLog incubator reader (Biolog, Inc., Hayward, CA).

2.6. Measurement of intercellular trehalose content

Cells of the test strains were inoculted into fresh improved LB-kanamycin broth to grow up to OD₆₀₀~0.3. pH of the mediums were regulated to 4.15 for 2 h and then the strains were incubated into fresh LB-kanamycin broth (pH ~7.0). When grown to an OD of 0.6–0.8, cells were collected by centrifugation and resuspended in lysis buffer (50 mM KH₂PO₄-K₂HPO₄, 500 mM NaCl, pH 6.0), and then by sonification and centrifugation at 12,000 g for 20 min at 4°C to remove cell debris. The amount of trehalose was measured using a HPLC (Dionex) system equipped with a Sepax HP-Amino column and a refractive index detector. The mobile phase was acetonitrile-water (75:25). The flow rate was 0.6 ml/min and column temperature was 30°C (Jiang et al., 2015). System was run in triplicate with replicates of the test strains.

3. Results

3.1. Isolation of acid-tolerance mutant

In order to isolate sigma D factor mutant strains with better acid tolerance, random mutagenesis based on the RAISE method was performed to construct *rpoD* mutant

libraries. A nearly 10^4 viable-mutant library was initially constructed and transformed into competent *E. coli* DH5 α cells. The transformed random mutagenesis libraries were then subjected to enrichment selection, that is, cells with improved acid-tolerance were gradually enriched through iterated inoculation of cell cultures into fresh medium with decreasing pH value. After four repeats of subculture, about 20 individual colonies were randomly picked from LB plates for plasmid extraction and their amino acid mutations in sigma D factor were confirmed by DNA sequencing. These plasmids were then reintroduced into fresh *E. coli* background to eliminate false positives. The mutant showing the greatest enhancement of acid-tolerance (as measured by the ratio of doubling time at different pH values tested) was used as a parent for the next generation of RAISE, with the selection cycle repeated for three generations (Supplementary Fig. S1 and Table 2). The isolated, characterized mutant (Mutant VII) at the third generation termed as the ‘best mutant’ showed improved overall growth at all pH values tested when compared to both the control strain as well as the best, isolated strains from the previous generations of selection (Fig. 1, Supplementary Fig. S2). As shown in Fig. 1, under the condition of neutral pH (pH 6.86), there is almost no difference in the growth of control strain and Mutant VII, with the growth rates both around 0.38 h^{-1} (Fig. 1A). However, when subjected to acid pH condition, cell growth of both the control strain and Mutant VII strain is inhibited, but Mutant VII demonstrates much better growth than the control strain (Fig. 1B, 1C and 1D). For instance, at pH 3.17, the growth rate of Mutant VII is 0.022 h^{-1} , while that of the control strain is 0.015 h^{-1} . Similar phenomenon

would be also observed in low-pH stress tolerant mutants of global regulator cAMP receptor protein (CRP) (Basak et al., 2014).

3.2. Transcriptional expression profiling of acid tolerance

In order to determine the transcriptional changes in cells harboring the mutant sigma D factor, the transcriptome of Mutant VII and the control strain were monitored using RNA-seq. Cells were first grown at neutral pH (in the absence of pH stress) to investigate on how RpoD factor mutation would affect their transcription profiles. A total of 273 genes (95 up-regulated and 178 down-regulated) were differentially expressed in Mutant VII by at least 2-fold compared to the control strain at P-value threshold of 0.05 (Fig. 2). These regulated genes were classified using the COG database. Generally, highly up-regulated genes were involved in carbohydrate metabolism (10 of 95 induced genes, 11%), regulatory functions and signal transduction (4%), amino acid biosynthesis (18%), nucleotide transport and metabolism (14%), the cell envelope (9%) and energy production and conversion (23%). Down-regulated genes were also involved in amino acid biosynthesis (17 of 178 repressed genes, 10%), energy production and conversion (15%), carbohydrate metabolism (14%), regulatory functions and signal transduction (17%), and nucleotide transport and metabolism (18%). Note that about 8% of the up-regulated and 14% of the down-regulated genes encoded proteins of hypothetical or unknown functions. Furthermore, no functional COG categories existed for 27 up-regulated genes and 38 down-regulated genes (Fig. 3,

Supplementary Table S1).

Specially, there are several typical up-regulated genes relevant to low pH stress (*gadABCX*, *ydeOP*, *yfdWX*, *emrKY*, *adiA*, *nrdH*, *slp*, *nanA* and *sdhA*) (2.72–19-fold, *p* value < 0.05), osmotic stress (*gmd*, *caiT*, *cpsB*, *otsAB*) (2.73–4.4-fold, *p* value < 0.05), oxidative stress (*hdeAB*, *katE*, *gshB* and *sodA*) (2.75–22-fold, *p* value < 0.05, also Supplementary Table S1). Moreover, RpoD factor altered the expression of many transcriptional regulators including the master regulators (ArcA, EvgA, H-NS and RpoS) and gene/operon-specific transcription factors (GadX, GadW, AppY, YdeO, KdgR). Among the listed regulators, 4 (*evgA*, *gadW*, *ydeO*, *arcA*, *gadX*) were induced, and another 4 (*rpoS*, *hns*, *appY*, *kdgR*) were repressed (Fig. 4, Supplementary Table S1).

3.3. qRT-PCR analysis for trehalose biosynthesis

Sigma D factor also induced *otsAB* operon (Fig. 4), encodes the two enzymes OtsA (trehalose-6-phosphate synthase) and OtsB (trehalose-6-phosphate phosphatase) which convert UDP-glucose to trehalose, resulting in the accumulation of trehalose within the cell. Thus the induction of the trehalose biosynthesis operon was expected to result in the accumulation of trehalose. As can be seen in Fig. 4, *otsA* and *otsB* genes had a 2.3-fold and 3.1-fold enhanced expression respectively in Mutant VII at pH 4.15 compared to the control based on the qRT-PCR analysis. This was further confirmed by HPLC analysis, that is, the Mutant VII strain accumulated 4.3-fold more trehalose than the control (0.64 ± 0.01 mg per 10^{11} cells instead of 0.15 ± 0.01 mg per 10^{11} cells).

3.4. Phenotypic profiling of acid tolerance

As shown in Fig. 5, using data from PM1, both Mutant VII and the control strain were metabolically versatile in their choice of carbon substrates. However, compared with the control strain, Mutant VII was found to be able to assimilate more efficiently 7 kinds of monosaccharide (e.g., L-arabinose, L-fucose, L-lyxose, D-galactose, D-ribose, α -D-glucose, D-psicose), 2 kinds of disaccharide (e.g., D-maltose, D-melibiose), as well as D-mannitol, D-galactonic acid- γ -lactone, L-asparagine, α -methyl-D-galactoside, and p-hydroxyphenyl acetic acid. While using PM10, we did find underlying phenotypic differences between Mutant VII and the control strain under acid stress conditions (pH 4.15). Compared with the control strain, Mutant VII could not only grow in the range of pH 4–10, but also make use of different kinds of amino acids under pH 4.5, including three kinds of neutral amino acid (L-alanine, glycine, L-valine), one acidic amino acid (L-glutamine), one basic amino acid (L-arginine), one aromatic amino acid (L-phenylalanine), one heterocyclic amino acid (L-Proline), as well as other amino acids (e.g., L-serine, L-threonine, L-methionine, hydroxy-L-Proline, L-homoserine, L-norvaline, α -amino-N-butyric acid) (Fig. 6).

4. Discussion

Sigma D factor in bacteria is responsible for the transcription of the majority of bacterial genes during exponential growth, commonly called the housekeeping gene. Housekeeping gene encodes products that are important for the major biosynthesis

1 pathways, including assimilation of amino acids and nucleotides, enzyme cofactors, cell
2 membrane and cell wall components, and carbon utilization enzymes in the glycolysis,
3 pentose phosphate shunt, and TCA cycle (Coleman et al., 2006). Herein, we employed a
4 random mutagenesis method named as RAISE to modify the *rpoD* gene with the global
5 disturbance of cell acid tolerance. The detailed protocol of RAISE is composed of only
6 three steps DNA fragmentation, attachment of a random short sequence and
7 reconstruction. The *rpoD* gene was first fragmented randomly by DNase I. Then,
8 several random nucleotides were attached to the 3' terminus of the fragment using TdT.
9 Each fragment with a tail of random nucleotides was finally reconstructed into a
10 full-length sequence by self-priming PCR. By adding this step, various lengths of
11 random insertions, deletions and substitutions would be succeeded in introducing into
12 *rpoD* gene (Fujii et al., 2006). Supplementary Fig. S1 identified the sequences of
13 mutants with the entire length of *rpoD* gene. Interestingly, five of the seven mutations
14 lead to the formation of a truncated RpoD factor, which is apparently beneficial to
15 increase overall acid tolerance. This deletion of several genes, arising from the RAISE
16 method, includes part of region 3 and the complete region 4 of the RpoD factor, which
17 is responsible for binding to the promoter region and to anti-sigma factors and further
18 impacting transcription (Gruber and Gross, 2003). Previous study deletion mutations in
19 the 603-607 site, have been involved in a reduction of transcriptional capacity at most
20 promoters tested and interactions with the alpha subunit of RNAP to interact with UP
21 regions in the DNA (Ross et al., 2003). The RAISE mutagenesis method suggest an

important distinction compared with the traditional protein-directed evolution, where mutations are typically additive in nature (Zhang et al., 1995).

From the growth data of different pH ranges, the Mutant VII obtained here showed the greatest improvement among the tested isolated strains (Table 2). When pH is around 4.2, the growth rate of Mutant VII is 0.034 h^{-1} , which is almost 2 times higher than the value of original strain used in this study (0.034 h^{-1} vs. 0.018 h^{-1}). Mutant VII has the best cell growth when challenged by stressor pH, which implies that RAISE technique can be employed to combine positive mutations to generate an even better mutant strain (Supplementary Fig. S2). Furthermore, as can be seen in Fig. 1, the performance difference between Mutant VII strain and control strain becomes more distinctive as the selection pressure increases. These findings suggest that mutations in sigma D factor may help mutant strain survive better than the control strain under stress. Since Mutant VII exhibits the best low pH tolerance among all mutants examined, it will be subjected to further omics investigation.

From the RNA-seq data, sigma D factor is able to bring about simultaneous changes referring to multiple metabolic functions including carbon metabolism, TCA cycle, nucleotide biosynthesis, amino acid utilization, as well as trehalose biosynthesis, except for the genes involved in acid resistance in *E. coli* (Fig. 3). A variety of structural as well as regulatory genes were up-regulated in the meantime to cope with the pH stressor in Mutant VII. For instance, *gadABC* (3.83–5.69-fold) and *adiA* (4.45-fold), encode glutamate and arginine decarboxylase system, which is proved to be a main part

of acid resistance mechanism in *E. coli* with help to maintain a near-neutral intracellular pH when cells are exposed to extremely acidic conditions (Ma et al., 2004). *caiT* (4.22-fold), encode L-carnitine/g-butyrobetaine antiport, was found to play a significant role in osmoregulation in *E. coli* (Zhang et al., 2012). While, *hdeA* (3.58-fold), encode protein HdeA with a chaperone-like function to prevent the aggregation of periplasmic proteins denatured under extremely acidic conditions (Gajiwala and Burley, 2000). The gene *hdeB* (22-fold), encode a structural homologue of HdeA protein, is located downstream of *hdeA* and predicted to form heterodimers with HdeA (Gajiwala and Burley, 2000). Furthermore, it is found in a previous study that overexpression of the response regulator EvgA induces 37 genes, including *gadABC* and *hdeAB* discussed just above, and confers acid resistance to exponentially growing *E. coli* (Masuda and Church, 2002). This acid resistance can be partly abolished by deletion of *ydeO* (3.21-fold) or *gadX* (2.83-fold), which are induced by EvgA overexpression, suggesting that the two genes are related to acid resistance in the log phase of the strain as well (Masuda and Church, 2002). Taken together the data from Supplementary Table S1, we have identified genes directly upregulated by EvgA and genes upregulated by EvgA-induced YdeO and GadX (Fig. 4). EvgA binds four consensus sequences and induces six operons (genes) including *ydeO* and *ydeP* (2.92-fold) (Masuda and Church, 2003). YdeO induces *Slp*, *HdeAB*, *GadX* and *GadABC*, which induce acid resistance by working co-operatively (Nishino et al., 2009). As can be seen in Fig. 4, many of the YdeO-activated genes are also activated by GadX, whereas *slp* is induced only by YdeO.

Slp is an outer membrane lipoprotein induced during carbon starvation and the late log phase (Stauffer and Stauffer, 2013). We suggest that RpoD factor may also control most genes indirectly through these regulators. For example, sigma D factor induces the regulator ArcA, one of the seven global regulators in *E. coli*, which then induces *yhiE* (Lin et al., 2013). Previous report has proved that overexpression of YhiE conferred acid resistance to exponentially growing cells of *E. coli* (Masuda and Church, 2003). Another example, since D competes for a fixed amount of RNA polymerase, the increase of RpoD may inevitably repress the expression of other sigma factors, e.g., RpoS (Fig. 4), which further leads to general stress resistance of cells, for instance, the Rcs phosphorelay, the PhoPQ two component system, and the (p)ppGpp-dependent stringent response, and the classic LexA-dependent SOS response (Hengge, 2009). However, the lower RpoS levels did not cripple the viability of *E. coli* within the external conditions (e.g., acid stress), which means that RpoD enables the growing cells to overcome external stresses much more effectively than originally thought (Battesti et al., 2011).

Initially, trehalose was thought to act solely as a reserve carbon source in a manner similar to that of starch, but plenty of emerging studies indicate that this non-reducing disaccharide instead has the important biological function of playing a major role in cell protection against various physical and chemical stresses, such as dehydration, desiccation, heat, cold, oxygen radicals, as well as acidic/alkali environmental conditions (Purvis et al., 2005; Jiang et al., 2013). From the qRT-PCR data of *otsA* and

1 *otsB* genes, it is suggested that the accumulation of trehalose could be attributed to the
2 *rpoD* expression in growing cells, which was consist with the previous reports
3 influenced by other regulators (Lin et al., 2013; Zhou et al., 2011). Furthermore,
4 consistent with such a role, the presence of free trehalose in apparently acid-stressed
5 cells could likewise have the same protective function, which was proved by other
6 groups (Ruhel et al., 2011). Previous research has revealed that accumulation of
7 constitutive trehalose in *E. coli* under stress conditions is thought to contribute to
8 maintenance of cell integrity and protein stability (Purvis et al., 2005).

9 Phenotype microarray (PM) technology was further used to investigate the
10 influence of sigma D factor on metabolic abilities of *E. coli*, which were scored using a
11 threshold calculation based on growth curve data previously described in Materials and
12 Methods section. From PM1 data, the change in mannitol utilization was particularly
13 interesting because the osmotolerance *E. coli* strain had an increased resistance to
14 osmotic shock caused by mannitol (Watanabe et al., 2003). The RNA-seq data showed
15 that the *mtl* operon, which belongs to the functional superfamily of the
16 phosphoenolpyruvate (PEP)-dependent sugar transporting phosphotransferases (PTS),
17 was up-regulated by RpoD factor. The *mtl* operon in *E. coli* consists of the *mtlA*, *mtlR*,
18 and *mtlD* genes that encode the mannitol transporter (enzyme IICBA^{mtl}), a
19 transcriptional regulator, and mannitol-1-phosphate dehydrogenase. *mtlA*, which is
20 located in the inner membrane, phosphorylates mannitol from the periplasm and
21 releases the phosphate ester into the cytoplasm. The mannitol-1-phosphate

1 dehydrogenase (MtlD) then oxidizes D-mannitol-1-phosphate to
2 D-fructose-6-phosphate. MtlD can also catalyze the reverse reaction and synthesize
3 mannitol. It has been reported that expression of the *E. coli mtlD* gene increased the
4 mannitol content and improved salt tolerance in bacteria (Hu et al., 2005). Contrary to
5 previous findings that the genes relevant to the uptake and initial degradation of carbon
6 sources decrease their expression to counter acid stress (Kwon et al., 2011), Mutant VII
7 displays up-regulation of many of these transporters in the neutral pH conditions (pH
8 7.0), for instance, maltose (*malEFG* operon), galactose (*galETKM* operon), melibiose
9 (*melR*), fucose (*fucPIKUR*), ribose (*rbsDACBK*), arabinose (*araC*) (Fig. 2), which
10 probably increases the intake of carbon and energy source to deal with the upcoming
11 environmental stress. This needs to be further investigated.

12 Form the PM10 data, it is showed that *astABCDE* operon, which is involved in
13 transport and degradation of L-arginine, is also up-regulated by RpoD factor under pH
14 4.5. The genes of this operon encode the enzymes of the arginine succinyltransferase
15 (AST) pathway and are required for growth of *E. coli* with arginine as the nitrogen
16 source. In *E. coli*, the well characterized arginine decarboxylase (ADC) pathway
17 accounts for only 3% of arginine catabolism, while the major contribution comes from
18 the ammonia-producing AST pathway (Shirai and Mizuguchi, 2003). Overproduction of
19 AST enzymes could protect the cells against damage caused by acidic environments
20 depends on the pH rise associated with ammonia production, even at pH values below
21 the minimal for growth or glycolysis. The PM data also show that *pheA*, which encodes

1 a chorismate mutase-prephenate dehydratase and is involved in the L-phenylalanine
2 catabolism (Polen et al., 2005), is induced by RpoD factor. Generally, all the above
3 genes involved in amino acid metabolism are induced by RpoD factor.

5 **5. Conclusions**

6 We have obtained an *E. coli* acid-tolerant variant through engineering its global
7 regulator RpoD factor. RAISE technique was employed to construct the random
8 mutagenesis libraries with over 10^4 cell population and Mutant VII was obtained
9 through multiwheel of acid shock screening. Mutant VII exhibited a significantly
10 improved growth rate as compared with the control strain at pH 3.17. We observed that
11 here is an overlap among low pH stress, oxidative stress and heat shock response.
12 Furthermore, there are a large number of host genes with a >2.0-fold change in
13 expression, including those responsible for nucleotides biosynthesis, carbon metabolism,
14 amino acid utilization, trehalose biosynthesis, except for acid resistance. To our
15 knowledge, this is the first report of modification and engineering of RpoD factor that
16 causes regulatory changes in unstressed *E. coli*. This current analysis is only a
17 beginning, and further research is needed to investigate the genome-wide transcriptional
18 response of the heterologous host *E. coli* expressing RpoD at various stresses.

20 **Acknowledgements**

21 We would like to thank Dr Rongrong Jiang from Nanyang Technological

University to give the pKSC plasmid. This work was supported by the National Basic Research Program of China (2012CB725204), the Natural Science Foundation of Jiangsu Province (BK20131406, BK20130917), the National Science Foundation for Young Scholars of China (21225626, 21406111, 21506101), the Six Talent Peaks Project in Jiangsu Province (2015-JY-009), and the Natural Science Foundation for Colleges and Universities in Jiangsu Province (14KJB530003).

References

- Abdel-Rahman, M.A., Tashiro, Y., et al., 2013. Recent advances in lactic acid production by microbial fermentation processes. *Biotechnol. Adv.* 31(6), 877-902.
- Alper, H., Moxley, J., et al., 2006. Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* 314(5805), 1565–1568.
- Alper, H., Stephanopoulos, G., 2007. Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab. Eng.* 9(3), 258-267.
- Basak, S., Geng, H., 2014. Rewiring global regulator cAMP receptor protein (CRP) to improve *E. coli* tolerance towards low pH. *J. Biotechnol.* 173, 68-75.
- Battesti, A., Majdalani, N., et al., 2011. The RpoS-mediated general stress response in *Escherichia coli*. *Annu. Rev. Microbiol.* 65, 189-213.
- Biondi, E.G., Tatti, E., et al., 2009. Metabolic capacity of *Sinorhizobium (Ensifer) meliloti* strains as determined by phenotype microarray analysis. *Appl. Environ. Microb.* 75(16), 5396-5404.

- 1 Chan, S., Kanchanatawee, S., et al., 2012. Production of succinic acid from sucrose and
2 sugarcane molasses by metabolically engineered *Escherichia coli*. *Bioresour.*
3 *Technol.* 103(1), 329-336.
- 4 Chun, A.Y., Liang, Y.X., et al., 2014. Elucidation of toxicity of organic acids inhibiting
5 growth of *Escherichia coli* W. *Biotechnol. Bioproc. E.* 19(5), 858-865.
- 6 Coleman, C., Baker, C., et al., 2006. The role of sigma factors in regulating bacterial
7 stress responses and pathogenesis. *Molecular Paradigms of Infectious Disease.*
8 Springer, US., pp. 438-501.
- 9 Dragosits, M., Mattanovich, D., 2013. Adaptive laboratory evolution-principles and
10 applications for biotechnology. *Microb. Cell Fact.* 12, 64.
- 11 Fujii, R., Kitaoka, M., et al., 2006. RAISE: a simple and novel method of generating
12 random insertion and deletion mutations. *Nucleic. Acids Res.* 34(4), e30.
- 13 Gajiwala, K.S., Burley, S.K., 2000. HDEA, a periplasmic protein that supports acid
14 resistance in pathogenic enteric bacteria. *J. Mol. Biol.* 295(3), 605–612.
- 15 Gruber, T.M., Gross, C.A., 2003. Multiple sigma subunits and the partitioning of
16 bacterial transcription space. *Annu. Rev. Microbiol.* 57(1), 441-466.
- 17 Hengge-Aronis, R., 2002. Signal transduction and regulatory mechanisms involved in
18 control of the σ^S (RpoS) subunit of RNA polymerase. *Microbiol. Mol. Biol. Rev.*
19 66(3), 373-395.
- 20 Hengge, R., 2009. Proteolysis of σ^S (RpoS) and the general stress response in
21 *Escherichia coli*. *Res. Microbiol.* 160(9), 667-676.

- Huang, C.F., Lin, T.H., et al., 2009. Enhanced ethanol production by fermentation of rice straw hydrolysate without detoxification using a newly adapted strain of *Pichia stipitis*. *Bioresour. Technol.* 100(17), 3914-3920.
- Hu, L., Lu, H., et al., 2005. Overexpression of *mtlD* gene in transgenic *Populus tomentosa* improves salt tolerance through accumulation of mannitol. *Tree Physiol.* 25(10), 1273.
- Ishihama, A., 2010. Prokaryotic genome regulation: Multi-factor promoters, multi-target regulators and hierarchic networks. *FEMS Microbial. Rev.* 34(5), 628–645.
- Jiang, L., Lin, M., et al., 2013. Identification and characterization of a novel trehalose synthase gene derived from saline-alkali soil metagenomes. *PloS One* 8 (10), e77437.
- Jiang, L., Cui, H.Y., Zhu L.Y., et al., 2013. Enhanced propionic acid production from whey lactose with immobilized *Propionibacterium acidipropionici* and role of trehalose synthesis in acid tolerance. *Green Chem.*, 17, 250–259.
- Kwak, S., Park, Y.C., et al., 2013. Biosynthesis of 3-hydroxypropionic acid from glycerol in recombinant *Escherichia coli* expressing *Lactobacillus brevis dhaB* and *dhaR* gene clusters and *E. coli* K-12 *aldH*. *Bioresour. Technol.* 135, 432-439.
- Kwon, Y.D., Kim, S., et al., 2011. Long-term continuous adaptation of *Escherichia coli* to high succinate stress and transcriptome analysis of the tolerant strain. *J. Biosci. Bioeng.* 111(1), 26-30.

- 1 Lin, Z., Zhang, Y., et al., 2013. Engineering of transcriptional regulators enhances
2 microbial stress tolerance. *Biotechnol. Adv.* 31(6), 986-991.
- 3 Marioni, J.C., Mason, C.E., et al., 2008. RNA-seq: an assessment of technical
4 reproducibility and comparison with gene expression arrays. *Genome Res.* 18(9),
5 1509-1517.
- 6 Masuda, N., Church, G.M., 2002. *Escherichia coli* gene expression responsive to levels
7 of response regulator EvgA. *J. Bacteriol.* 184, 6225–6234.
- 8 Masuda, N., Church, G.M., 2003. Regulatory network of acid resistance genes in
9 *Escherichia coli*. *Mol. Microbiol.* 48(3), 699-712.
- 10 Mazumdar, S., Blankschien, M.D., et al., 2013. Efficient synthesis of L-lactic acid from
11 glycerol by metabolically engineered *Escherichia coli*. *Microb. Cell Fact.* 12(7),
12 10-1186.
- 13 Nishino, K., Senda, Y., et al., 2009. Role of the AraC–XylS family regulator YdeO in
14 multi-drug resistance of *Escherichia coli*. *J. Antibiot.* 62(5), 251-257.
- 15 Polen, T., Krämer, M., et al., 2005. The global gene expression response of *Escherichia*
16 *coli* to L-phenylalanine. *J. Biotechnol.* 115(3), 221-237.
- 17 Purvis, J.E., Yomano, L.P., et al., 2005. Enhanced trehalose production improves
18 growth of *Escherichia coli* under osmotic stress. *Appl. Environ. Microb.* 71(7),
19 3761-3769.
- 20 Richard, H., Foster, J.W., 2004. *Escherichia coli* glutamate- and arginine-dependent
21 acid resistance systems increase internal pH and reverse transmembrane potential.

1 J. Bacteriol. 186(18), 6032-6041.

2 Ross, W., Schneider, D.A., et al., 2003. An intersubunit contact stimulating
3 transcription initiation by *E. coli* RNA polymerase: interaction of the α C-terminal
4 domain and ς region 4. *Gene Dev.* 17(10), 1293-1307.

5 Rowbury, R.J., Goodson, M., et al., 1994. Sodium chloride induces an
6 NhaA/NhaR-independent acid sensitivity at neutral external pH in *Escherichia coli*.
7 *Appl. Environ. Microbiol.* 60 (5), 1630–1634.

8 Ruhal, R., Aggarwal, S., et al., 2011. Suitability of crude glycerol obtained from
9 biodiesel waste for the production of trehalose and propionic acid. *Green Chem.*
10 13(12), 3492-3498.

11 Shirai, H., Mizuguchi, K., 2003. Prediction of the structure and function of AstA and
12 AstB, the first two enzymes of the arginine succinyltransferase pathway of arginine
13 catabolism. *FEBS Lett.* 555(3), 505-510.

14 Stauffer, L.T., Stauffer, G.V., 2013. Multiple roles for the sRNA GcvB in the regulation
15 of Slp levels in *Escherichia coli*. *ISRN Bacteriol.* 2013.

16 Watanabe, S., Hamano, M., et al., 2003. Mannitol-1-phosphate dehydrogenase (MtlD) is
17 required for mannitol and glucitol assimilation in *Bacillus subtilis*: possible
18 cooperation of *mtl* and *gut* operons. *J. Bacteriol.* 185(16), 4816-4824.

19 Zhang, H., Chong, H., et al., 2012. Random mutagenesis of global transcription factor
20 cAMP receptor protein for improved osmotolerance. *Biotech. Bioeng.* 109(5),
21 1165-1172.

- 1 Zhang, X., Baase, W., et al., 1995. Enhancement of protein stability by the combination
2 of pointmutations in T4 lysozyme is additive. *Protein Eng.* 8(10), 1017-1022.
- 3 Zhou, Z., Zhang, W., et al., 2011. Genome-wide transcriptome and proteome analysis of
4 *Escherichia coli* expressing IrrE, a global regulator of *Deinococcus radiodurans*.
5 *Mol. BioSyst.* 7(5), 1613-1620.

1

Table 1. Primers used in this study

Designation	Oligonucleotide sequence (5'–3')	Digestion enzyme
<i>rpoD</i> -sense-EcoRI	CCGAATTC TGATTTAACGGCTTAAGTGCCGAAG	<i>EcoRI</i>
<i>rpoD</i> -anti-HindIII	TAAAGCTT GCCGGGTCGGCGGTAAC	<i>HindIII</i>
sence-M13	CGCCAGGGTTTTCCCAGTCACGAC	
anti-M13	AGCGGATAACAATTCACACAGGA	
<i>otsA</i> -sense-EcoRI	CCGAATTC ATGAGTCGTTTAGTCGTAGTATCTA	<i>EcoRI</i>
<i>otsA</i> -anti-HindIII	TAAAGCTT CTACGCAAGCTTTGGAAAGGTAGCA	<i>HindIII</i>
<i>otsB</i> -sense-EcoRI	CCGAATTC GTGACAGAACCGTTAACCGAAACCC	<i>EcoRI</i>
<i>otsB</i> -anti-HindIII	TAAAGCTT TTAGATACTACGACTAAACGACTCA	<i>HindIII</i>

2

1 **Table 2.** Enhancement of acid tolerance through engineered sigma D factor.

	Doubling Time (h)	Ratio of doubling times ($t_{d,control} / t_{d,engineered\ mutant}$)		
pH	control	Round 1	Round 2	Round 3
6.86	0.78	1.01	0.82	0.77
4.15	33.0	1.32	18.9	20.4
3.62	46.2	1.65	30.1	31.5
3.17	40.8	1.94	35.8	36.5

2 Enhancements in the fold reduction of doubling time are presented for decreasing pH
3 for the isolated mutant in each of the three rounds of directed evolution.
4 Measurements of growth rate were conducted under the above pHs and the fold
5 reduction in doubling times were calculated as a ratio of the doubling time of the
6 control to that of the mutant strain for each specific pH. The control strain is an *E. coli*
7 DH5 α strain harboring the wild-type sigma D factor on the pKSC plasmid (see
8 Materials and Methods). The mutants from both rounds 2 and 3 show significant
9 increases in the growth rate at lower pH.

10

1
2 **Fig. 1.** Growth profiles of the control strain and Mutant VII under various pHs: (A)
3 pH 6.86, (B) pH 4.15, (C) pH 3.62, (D) pH 3.17. Each overnight colony (both mutant
4 and control) was inoculated into fresh improved LB-kanamycin broth (5 g/l NaCl) at
5 37°C with 200 rpm. All experiments were performed in three replicates.

Figure 2

1 **Fig. 2.** Selected differentially expressed genes grouped by GO functional
2 classification.

Figure 3

1
2 **Fig. 3.** Up- and down-regulation of the metabolic pathway and the transport system genes *E. coli* expressing *rpoD*. The colour indicates
3 significant changes ($P > 0.05$) in gene expression or genes and pathways. Red, overexpression; blue, reduced expression; black, change <
4 1.8-fold. Dashed arrows indicate multiple enzymatic steps. Abbreviations: UMP, uridine 5'-monophosphate; PRPP, Phosphoribosyl
5 pyrophosphate; PRA, phosphoribosylamine; GAR, glycinamide ribonucleotide; FGAR, N-formylglycinamide ribonucleotide; FGAM,
6 N-formylglycinamidine ribonucleotide; AIR, 5-aminoimidazole ribonucleotide; CAIR, carboxy aminoimidazole ribonucleotide; SAICAR,

- 1 N-succinyl-5-aminoimidazole-4-carboxamide ribonucleotide; AICAR, 5-aminoimidazole-4-carboxamide ribonucleotide; FAICAR,
- 2 5-formaminoimidazole-4-carboxamide ribonucleotide; IMP, inosine monophosphate; PEP, phosphoenolpyruvic acid; OAA, oxaloacetic acid.
- 3 The pathway map was adapted from the KEGG pathway database.
- 4

Figure 4

1

2 **Fig. 4.** RpoD regulation in *E. coli*. The transcriptional regulatory network, an overlap
3 of strain response towards various stresses, was constructed from cDNA data. Solid
4 lines, proven regulatory interactions; broken lines, predicted interactions. Oval shapes,
5 *E. coli* inducers; squares, repressors. Italic, regulated genes.

Figure 5

Fig. 5. The mRNA expression (normalized to naive spinal cord) for trehalose-6-phosphate synthase (*otsA*), trehalose-6-phosphate phosphatase (*otsB*) by qPCR at pH 4.15, OD₆₀₀~0.6.

Figure 6

Fig. 6. Phenotypic changes in PM1 assays. Significant changes are enclosed in boxes and indicated by arrows. Yellow indicates that growth of the control strain and growth of the Mutant VII were similar. Red indicates faster growth of the control strain. Green indicates faster growth of the Mutant VII. Comparison of the control strain and Mutant VII: well A02, L-arabinose; well A06, D-galactose; well B04, L-fucose; well B10, D-mannitol; well C02, D-galactonic acid-g-lactone; well C04, D-ribose; well C09, a-D-glucose; well C10, D-maltose; well C11, D-melibiose; well D01, L-asparagine; well D08, a-Methyl-D-Galactoside; well H02, p-hydroxyphenyl acetic acid; well H05, D-psicose; well H06, L-lyxose.

Figure 7