IDENTIFICATION OF PREFERENTIALLY EXPRESSED GENES IN *ESCHERICHIA COLI* CULTURES IN LATE STATIONARY PHASE

IDENTIFICATION OF PREFERENTIALLY EXPRESSED GENES IN *ESCHERICHIA COLI* CULTURES IN LATE STATIONARY PHASE

By

LINGZI KONG, B, SC

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

McMaster University MASTER OF SCIENCE (2013) Hamilton, Ontario (Biology)

TITLE: Identification of preferentially expressed genes in *Escherichia coli* cultures in late stationary phase

AUTHOR: Lingzi Kong, B. Sc

SUPERVISOR: Professor H.E. Schellhorn

NUMBER OF

PAGES: x, 105

Abstract

Escherichia coli cells undergo many morphological and physiological changes to survive in late stationary phase due, in part, to the exhaustion of nutrients and accumulation of inhibitory metabolites. Previous microarray data in our lab provided the general profile of gene expression from exponential phase to 48 h of incubation. The goal of this study is to use qPCR to validate the preferentially expressed genes in late stationary phase determined by the microarray data. The expression of three genes (*hha*, *tomB* and *emrK*) with increased expression levels from 24 h to 48 h of incubation and another two RpoS-dependent genes (*bolA* and *osmY*) with decreased expression from 24 h to 48 h of incubation were chosen as the target genes. RNA was first extracted from exponential phase, early stationary phase, 24 h and 48 h bacterial cultures. Hot phenol-chloroform and a commercial kit were used to isolate RNA. All these methods recovered low RNA in late stationary phase. qPCR partly confirmed the previous microarray data. The expression of *hha*, *tomB* and *emrK* was validated to increase from 24 h to 48 h of incubation. The discordant results between qPCR and microarray data may be due to the low transcript abundance of target genes and genomic DNA contamination in RNA extracted from 48 h bacterial cultures.

Acknowledgement

I would like to thank my supervisor, Dr. Herb E. Schellhorn for his supervision and giving me the opportunity to work in this great lab. I would like to thank Dr. Gupta, my co-supervisor, and Dr. Kolasa, my committee chair, for critically reviewing this thesis and providing valuable suggestions.

I would like to thank all the members of the Schellhorn lab: Dr. Sarah Chiang, Shirley Wong, Rosemary Yu, Steven Botts, Mohammad Howard-Azzeh, Mohammad Mohiuddin, Andrew O'Brien, Myra Sohail, Quan Yao, Danya Thayaparan, Pardis Haghighi ,Wenjie Zheng and Xiaoli Zhao. I am grateful for the help and wonderful friendship from you. I would also thank Dr. Tao Dong, whose previous work provided a solid foundation and reference for me.

I would also thank my parents for their support and my friends in Canada and back in China. They patiently listened to me when I needed emotional support, providing helpful suggestions and encouragements.

Table of Content

List of Figures

List of Tables

Chapter 1. Introduction

1.1 The *Escherichia coli* life cycle

Escherichia coli is a common intestinal bacteria in humans and warm blooded animals (Hartley et al. 1977) that can survive in nature for extended periods of time. The *E. coli* life cycle consists of five stages, determined by measuring cells' viability grown in the rich culture at 37˚C (Finkel 2006): lag phase, exponential phase, stationary phase, death phase and long-term stationary phase.

In lag phase, cells process several metabolic activities to thrive in new nutritional environment, which include repairing damage from last stationary phase and synthesizing necessary cellular components for growth (Rolfe et al. 2012). In exponential phase, bacteria grow very fast and cells are generally at a low density. The consumption of nutrients is low and the synthesized metabolites cannot affect growth (Sezonov et al. 2007). Due to the exhaustion of nutrients and the accumulation of metabolite wastes, bacteria enter stationary phase. After the transition from exponential phase to stationary phase, *E. coli* cells undergo many morphological and physiological changes (discussed in detail in section 1.2). Then, cells enter a death phase where most cells lose viability. *E. coli* cells enter death phase in LB media after 3 days of continuous incubation (Finkel 2006). Though most cells die, the surviving cells can survive for months or years, which is called long-term stationary phase (Finkel 2006).

1.2 Morphological changes of *E. coli* in stationary phase

E. coli cells in stationary phase become much smaller and almost spherical in contrast to the rod shape of the exponential phase cells (Leland et al. 1983). The reductive division and dwarfing can account for this phenomenon (Nystrom 2004). The expression of *bolA* increases in early stationary phase, which produces the round-shaped cells (Santos et al. 2002). Reductive division is a phenomenon when cells complete DNA replication and division but do not grow anymore (Nystrom 2004). Dwarfing is a form of self-digestion and can be triggered by starvation. Dwarfing causes endogenous material degradation, including the cytoplasmic membrane and cell wall (Nystrom 2004).

The cell envelope properties of stationary phase *E. coli* are different from the exponential phase cells, including the properties of the outer membrane, periplasm and the inner membrane (Rutishauser et al. 1996). For example, in the outer membrane, the concentration of lipopolysaccharides increases (Navarro Llorens et al. 2010) and the total amount of protein decreases (Allen and Scott 1979). The thickness of peptidoglycan layer increases, which can help cells bear external stress (Mengin-Lecreulx and van Heijenoort 1985). In the inner membrane, the unsaturated membrane fatty acids are converted to the cyclopropyl derivatives (Cronan 1968). As a result, the inner membranes become less fluid and permeable (Kolter et al. 1993). All these changes in the cell envelope can help cells protect against external stress (Kolter et al. 1993). In the cells, genomic DNA is packed to form nucleoid (Robinow and Kellenberger 1994), which can protect DNA against the environmental stresses (Ohniwa et al. 2006). This phenomenon is fulfilled by

Dps (DNA-binding protein from starved cells). Dps is highly abundant in stationary phase *E. coli*, which can form a highly ordered and stable nucleoprotein complex, resulting in the condensed nucleoid (Wolf et al. 1999). The chromosomes undergo some topological changes in stationary phase cells. *E. coli* begins to have the plasmids with negative superhelical density after several hours in stationary phase (Dorman et al. 1988; Reyes-Dominguez et al. 2003). The changes in supercoiling affect promoter function and initiate transcription (Dorman et al. 1988). In addition, the number of chromosomes in stationary phase *E. coli* is different. For instance, in rich medium, most cells contain 2 or 4 chromosomes while in minimal medium, cells contain 1 or 2 chromosomes (Akerlund et al. 1995). After entering stationary phase, RNA stability decreases. 20% to 40% of total RNA is lost in the first a few hours after entering starvation conditions (Mandelstam 1960). This degradation during starvation is mostly due to rRNA degradation. However, ribosomal proteins remain relatively stable, which are important for starved cells to recover more rapidly (Deutscher 2003). The degradation of ribosomal RNA during starvation is rapid and extensive. RNA degradation begins as soon as the nutrients become limited and even before growth ceases and cells die (Davis et al. 1986). The degradation of ribosomal RNA starts with endonucleolytic cleavages such as RNase I, which is then followed by exoribonuclease cleavages (Deutscher 2003). The changes in cell membrane at the onset of stationary phase can help RNase I in the periplasmic space enter the cytoplasm and then cause ribosomal RNA degradation (Deutscher 2003).

1.3 Physiological changes of *E. coli* in stationary phase

In stationary phase, *E. coli* cells undergo many physiological changes to adapt to starvation and stress conditions. For instance, the rate of protein turnover increases, which is due to the synthesis of peptidase and proteases (Groat et al. 1986). In addition, *E. coli* cells show more resistance to external stress, such as heat shock, oxidative stress (Jenkins et al. 1988) and osmotic stress (Jenkins et al. 1990) than exponential phase cells. In cells, several regulators play important roles in those physiological changes.

1.3.1 Important regulators in *E. coli* stationary phase

Bacterial regulators can be categorized into two groups: specific regulators and general regulators. The specific regulators, such as FadR for regulating *fad* operons to digest endogenous membrane constituents and provide carbon and energy for cells to grow (Farewell et al. 1996), are induced under particular stress conditions and regulate genes in response to specific stress. The general regulators, such as ppGpp and RpoS, can control the expression of a large group of genes in response to multiple environmental stresses. Some important general regulators in stationary phase are introduced below.

1.3.1.1 ppGpp

ppGpp is a key regulator in bacteria. ppGpp can rapidly mediate gene expression in response to diverse stresses, such as amino acid starvation (Haseltine and Block 1973), fatty acid starvation (Battesti and Bouveret 2006) and iron limitation (Vinella et al. 2005). This regulation is called the stringent response. ppGpp can decrease rRNA biosynthesis

and DNA replication. ppGpp increases the expression of *rpoS*, the biosynthesis of stress protein (Magnusson et al. 2005). RelA and SpoT can mediate the concentration of ppGpp in cells. RelA synthesizes ppGpp while SpoT can both produce and degrade it (Xiao et al. 1991). Amino acid starvation induces the expression of *relA* (Xiao et al. 1991), while the limitation of carbon source (Xiao et al. 1991), phosphorous (Spira et al. 1995), iron (Vinella et al. 2005) and fatty acid (Seyfzadeh et al. 1993) increases the expression of *spoT*. ppGpp can also bind to RNA polymerase and regulate the expression of many genes (Barker et al. 2001). In addition, ppGpp positively regulates the expression of *rpoS*, another global regulator in stationary phase (Gentry et al. 1993).

1.3.1.2 RpoS and sigma factor competition

RpoS is a bacterial transcription factor that can bind RNA polymerase to gene promoters and initiates gene transcription. Different sigma factors bind to core polymerase under specific conditions. The composition of sigma factors varies in different bacteria. In *E. coli*, in addition to RpoS, there are six other sigma factors (Merrick 1993). Sigma factors can be clustered into two families: sigma 70 family and sigma 54 family based on the structural differences (Paget and Helmann 2003). The sigma 70 family can be further divided into 4 groups. Group 1, which includes RpoD, can bind to the promoters of most housekeeping genes (Ishihama 2000). RpoD is the most abundant sigma factor in *E. coli* (Jishage et al. 1996). RpoS belongs to Group 2 of the sigma factors. Although RpoS is considered as a global regulator responsible for the expression of many genes under stress conditions and in stationary phase (Schellhorn

1995), RpoS also mediates gene expression in exponential phase (Dong et al. 2008). Group 3 comprises of RpoH and FliA (Gruber and Gross 2003). RpoH is induced under heat shock (Grossman et al. 1984) and FliA controls flagellar biosynthesis (Helmann 1991). RpoE and FecI are categorized into Group 4 (Gruber and Gross 2003). RpoE is responsible for gene regulation in response to the heat shock and hyperosmotic shock (Bianchi and Baneyx 1999). RpoE is also involved in the repair of misfolded protein (Mecsas et al. 1993). FecI can regulate the expression of genes in response to iron deprivation in the presence of ferric citrate (Braun et al. 2003). Group 4 sigma factors are typically induced in response to the extracellular stress (Gruber and Gross 2003). RpoN is structurally distinct from sigma 70 family (Lonetto et al. 1992). RpoN controls the expression of genes in response to nitrogen limitation (Hirschman et al. 1985), heat shock (Kuczynska-Wisnik et al. 2010) and flagellar biosynthesis (Zhao et al. 2010).

RpoS is considered to be the second vegetative sigma factor in non-optimal conditions (Weber et al. 2005). Sigma factors compete with each other due to the limited amount of RNA polymerase. The overproduction of RpoD can decrease the expression of RpoS regulon and the cells become stress sensitive (Farewell et al. 1998). Also, the overproduction of RpoS reduces the expression of genes requiring RpoD (Farewell et al. 1998). This competition is highly dependent on the nutritional environment as well as the concentration of ppGpp (Jishage et al. 2002). ppGpp can induce the expression of *rpoS* under stress and starved conditions (Jishage et al. 2002). RpoS can positively regulate the expression of genes in response to many stresses, such as *katE* for oxidative stress (Schellhorn and Hassan 1988) and *ostB* for temperature shift (Hengge-Aronis et al. 1991).

RpoS also represses the expression of many genes, such as the genes responsible for flagellar biosynthesis and TCA cycle (Patten et al. 2004). Genes in TCA cycle are important for cell metabolism and the production of energy (Patten et al. 2004). The repression those genes by RpoS may be important for long-term survival, because strains with higher RpoS levels are more resistant to environmental stress but have limited metabolizing capabilities. On the other hand, strains with lower RpoS levels can better survive in low nutrient environments, but also show lower resistance to external stress. This trade-off phenomenon is called self-preservation and nutritional competence (SPANC) (Ferenci 2005). In *E. coli*, the amount of RpoS is far less than RpoD (Jishage et al. 1996) and RpoS has the weakest affinity to core RNA polymerase compared with other sigma factors (Maeda et al. 2000). In stationary phase, except for the increased expression level of *rpoS*, other transcription factors also play important roles to help RpoS compete with RpoD. For example, the anti-sigma factor Rsd is induced to sequester RpoD in stationary phase, allowing more RpoS to bind to RNA polymerase (RNAP) (Jishage and Ishihama 1998). ppGpp also positively regulates the expression of *rpoS* (Jishage et al. 2002). The composition and expression of the RpoS regulon is modified in response to different environmental conditions. Almost 300 genes are regulated by RpoS in exponential phase (Dong et al. 2008). Among the highly RpoS-dependent genes, only 12 of 75 are RpoS-dependent in both stationary and exponential phases (Dong et al. 2008).

1.3.1.3 Lrp

The leucine-responsive regulatory protein (Lrp) regulates more than 400 genes. Most of them are induced in stationary phase (Tani et al. 2002). Those genes are involved

in the response to nutrient limitation, acid and osmotic stress (Tani et al. 2002). Lrp also plays an important role in amino acid biosynthesis (Ernsting et al. 1992) and catabolism, nutrient transport, pili synthesis and other cellular functions (Brinkman et al. 2003).

1.3.1.4 IHF

IHF (integration host factor), is a histone-like protein, which can bind to DNA sequence and mediate gene expression (Mangan et al. 2006). IHF is highly abundant in *E. coli* and shows a growth phase-dependent expression (Ditto et al. 1994). IHF regulates the expression of many genes, such as *osmY* for the response to oxidative stress (Colland et al. 2000) and *dps* for curli production (Altuvia et al. 1994). IHF can also affect DNA replication and site-specific recombination (Dhavan et al. 2002).

1.3.2 Metabolic activities

As the cells become starved, the overall metabolic rate and ATP levels decrease (Chapman et al. 1971). However, starved *E. coli* cells are not truly dormant, which is in contrast with the dormant state of the spores in Gram-positive strains (Kolter et al. 1993). The rate of protein synthesis decreased to 20% of that in exponential phase in the first hour of starvation and then remains roughly at a similar level for at least 37 h (Reeve et al. 1984). In starvation conditions, the metabolisms of carbohydrates, amino acids and phospholipids decrease (Tani et al. 2002). Dimerization of ribosomes may inhibit protein translation (Wada et al. 1995), which further affects the metabolism. The ribosome (70 S) in *E. coli* consists of two subunits: a small (30 S) and a large (50 S) subunit. The small

subunit has a 16 S RNA subunit and the large one has a 5 S subunit and a 23 S subunit. In stationary phase, the 70 S subunit can be dimerized to 100 S, leading to the inactivity of the protein translation (Wada et al. 1995).This dimerization is regulated by the ribosome modulation factor (RMF), which is induced in the stationary phase. RMF can bind to 50 S ribosomal subunits, resulting in the dimerization of 70 S ribosomes. ppGpp, an important global regulator in stationary phase, can positively control the expression of *rmf* (Izutsu et al. 2001). The decreased metabolism in TCA cycle (Nystrom 1994) is regulated by AcrA, which reduces the production of NADH/$FADH_{2.}$ This regulation can repress the respiratory activity (Nystrom et al. 1996), which helps cells resist the damage of reactive oxygen species produced by respiration (Apel and Hirt 2004).

Dwarfing is a form of self-digestion and is similar to autophagy in eukaryotes (Nystrom 2004). Autophagy allows recycling of unnecessary cellular components to ensure cells to survive during starvation. *fad* operons digest endogenous membrane constituents and use long-chain fatty acids as a carbon source to provide cells with carbon and energy (Farewell et al. 1996). Apart from the increased expression of *fad* genes, peptidase-dependent autophagy is also important for de novo protein synthesis by providing enough amino acids in the absence of exogenous carbon source (Groat et al. 1986). In starved cells, protein turnover increases 5-fold in the early stages of starvation. Cells without ClpAP or ClpXP proteases exhibit an accelerated die-off in late stationary phase (Weichart et al. 2003).

Some secondary metabolites, such as antibiotics, are synthesized in stationary phase. For instance, microcins are produced to fight against closely related bacteria (Duquesne et al. 2007). QS (quorum sensing) molecules are also produced at the onset of stationary phase (Keller and Surette 2006). Bacteria release signaling molecules into environment. When the concentration reaches a threshold, cells will sense it and then respond to it (Keller and Surette 2006). QS can also control bioluminescence, the expression of virulence genes, biofilm formation and the conjugal transfer of plasmid DNA (Lazazzera 2000). In *E. coli*, the multidrug transporters can exude the QS molecules outside the membrane (Yang et al. 2005). Autoinducer-2 is one kind of QS molecules in *E. coli* (Surette and Bassler 1998).

1.3.3 Programmed cell death

Bacterial cell death in stationary phase can be caused by stochastic cellular death or programmed cell death (PCD). The accumulation of damaged molecules due to oxidative damage will lead to bacterial death (Dukan and Nystrom 1998). Programmed cell death can limit the spread of viral infection and provide nutrients to other cells (Engelberg-Kulka et al. 2006). In addition, programmed cell death decreases the mutation rate by eliminating cells with damaged DNA (Lewis 2000). The programmed cell death is mediated by toxin-antitoxin (TA) modules. For instance, starvation could induce the expression of *relBE* operon (Christensen et al. 2001). Oxidative stress could increase the expression of *hipBA* (Kawano et al. 2009). Another widely studied TA module in *E. coli* is *mazEF* operon. *mazE* encodes antitoxin and *mazF* encodes toxin. The increased

concentration of ppGpp can repress the expression of *mazEF* (Marianovsky et al. 2001) . The toxin is more stable than the antitoxin (Engelberg-Kulka et al. 2006). Therefore, ppGpp can accumulate the toxin. However, other research indicates that the toxin can only inhibit mRNA translation instead of killing the bacteria (Buts et al. 2005). Again, this phenomenon confirms the important role of ppGpp in stationary phase cells.

1.3.4 Resistance to external stress

E. coli cells in stationary phase show more resistance to external stress, such as heat shock (Jenkins et al. 1988), oxidative stress and osmotic stress than exponential phase cells (Lange and Hengge-Aronis 1991).

1.3.4.1 Oxidative stress response

Bacteria face oxidative stress under aerobic condition due to the production of reactive oxygen species (ROS), including superoxide radicals (O_2^-), hydrogen peroxide (H_2O_2) and reactive nitric oxide (NO) (Chiang and Schellhorn 2012). Bacteria develop specific mechanisms to resist the oxidative stress and ensure cells to survive in aerobic conditions. Two catalases, HPI and HPII encoded by *katE* and *katG* are induced in stationary phase and under oxidative stress (Schellhorn 1995). The expression of *katG* is regulated by OxyR, an important factor in oxidative stress response (Gonzalez-Flecha and Demple 1997). KatE is the major catalase in stationary phase and is highly RpoS-dependent (Schellhorn and Hassan 1988). Anther mechanism is the production of

superoxide dismutase (SOD). *E. coli* has three superoxide dismutase, encoded by *sodA*, *sodB* and *sodC* (Imlay and Imlay 1996).

1.3.4.2 Acid response

E. coli has three acid resistance systems. There are RpoS-dependent oxidative AR system, arginine-dependent decarboxylase AR system and glutamate-dependent decarboxylase AR system (Castanie-Cornet et al. 1999). RpoS plays an important role in those AR systems, especially the RpoS-dependent oxidative system (Lin et al. 1996). In the arginine-dependent system, *adiA*, encoding arginine decarboxylase is positively regulated by RpoS (Vijayakumar et al. 2004). In the glutamate AR system, two glutamate decarboxylase genes, *gadA* and *gadB* are both induced in stationary phase via the regulation of RpoS (Castanie-Cornet et al. 1999).

1.3.5 High genomic mutation rate in stationary phase

The genomic mutation rate in stationary phase *E. coli* cells is high (Loewe et al. 2003). The accumulated mutation is ascribed to the inadequate repair system. The induction of error-prone DNA polymerases increases the amount of DNA lesions that cannot be totally repaired by the DNA repair system (Saint-Ruf and Matic 2006). In addition, the repression of methyl-directed mismatch repair (MMR) also contributes to the accumulated mutations in *E. coli* (Saint-Ruf and Matic 2006). They are both under control of RpoS. In stationary phase, RpoS represses the expression of *mutS* and *mutH*, which play important roles in DNA replication and post-replicative error repair in MMR

system. *dinB*, encoding the error-prone DNA polymerase (PolIV) is induced under control of RpoS. polIV is the main factors for the adaptive mutations (Saint-Ruf and Matic 2006). It is possible that the accumulated mutations help cells adapt to the changing environment (Hall 1991). Natural selection selects the fittest mutant for a specific environmental condition from these stress-induced mutations (Loewe et al. 2003). However, more researches are still needed to elucidate the role of stress-induced mutations.

1.4 Preliminary microarray data and rationale of this study

Most research focuses on the gene expression in exponential phase and the early stationary phase. The literature about gene expression in late stationary phase is limited. In this project, we want to elucidate gene expression in late stationary phase, which will help better understand how *E. coli* adapts to the changing environment.

Microarray data of *E. coli* MG1655 gene expression in exponential phase (OD_{600}=0.3) and early stationary phase (OD_{600}=1.5) were obtained from Dong (Dong et al. 2008). Another dataset of *E. coli* MG1655 gene expression at 24 h and 48 h of incubation were obtained from S. Chiang (unpublished data). These data provided general profiles of gene expression over time. Different genes showed different expression level changes from 24 h to 48 h of incubation. For instance, the expression levels of *hha* and *tomB* increased dramatically from 24 h to 48 h of incubation. The expression of genes encoding rRNA, such as *rrsA*, maintained at the same level. Some gene expression levels decreased dramatically from 24 h to 48 h of incubation, such as *osmB*, which encodes lipoprotein

and is positively regulated by RpoS in stationary phase (Patten et al. 2004). Some gene expression levels decreased slightly from 24 h to 48 h of incubation, such as *rpoS*, encoding a transcriptional factor responsible for the regulation of many genes in stationary phase and under stress conditions. In this study, we are interested in the genes with increased expression levels from 24 h to 48 h of incubation, as those gene functions may help *E. coli* adapt to the late stationary phase, such as resisting external stress, avoiding autolysis and decreasing some metabolic activities.

Based on S. Chiang's microarray data, *hha, tomB, emrK, yafY, ytfE, mcrC, yfdK* and *yhcE* are the top 8 genes with the highest increase in expression levels between 24 h and 48 h of incubation after normalization. In addition, *bolA* and *osmY* are two RpoS-dependent genes whose expression levels increase when cells enter into stationary phase were used as controls to confirm the reliability of the qPCR results. The expression of *bolA* and *osmY* decreased from 24 h to 48 h of incubation determined by microarray results.The functions and gene transcript abundances at different growth phases are discussed below.

1.4.1 *hha*

Hha (high hemolysin activity modulator) belongs to Hha-YmoA family which can regulate many gene expressions in gram-negative bacteria (Madrid et al. 2007). Hha reduces hemolysin production by repressing the expression of *hly*. This repression of *hly* is fulfilled by the H-NS-Hha complex. Hha can interact with the nucleoid-associated protein H-NS to form a hetero-oligomeric complex, which improves H-NS repression

efficiency, especially at low temperatures (Nieto et al. 2000). H-NS is a major component of the bacterial nucleoid protein and negatively regulates the expression of many genes (Hommais et al. 2001). Apart from *hly*, Hha also represses the expression of virulence genes in pathogenic *E. coli*, such as genes at the locus of enterocyte effacement (LEE island) (Sharma and Zuerner 2004). H-NS Hha complex can also specifically silence the genes acquired by horizontal gene transfer (HGT) (Banos et al. 2009). To efficiently silence the expression of HGT DNA, H-NS recruits Hha to regulate HGT DNA while the genomic DNA is regulated by H-NS alone (Banos et al. 2009). Hha is highly induced in *E. coli* biofilms compared with planktonic cells (Ren et al. 2004). Hha decreases the initial biofilm formation not only by repressing the transcription of rare codon tRNA , such as *argU*, *ileXY* and *proL,* thus to regulate fimbriae production at translational level, but also directly repressing the transcription of fimbrial genes, such as *ihfA* and *fimA* (Garcia-Contreras et al. 2008). Hha causes cell lysis and biofilm dispersal by activating prophage lytic genes, such as *rzpD, yfiZ* and *appY,* and inducing the production of ClpP/ClpX proteases (Garcia-Contreras et al. 2008). In some *E. coli* strains, nutrient starvation is an environment stimuli for the release of microbes from a biofilm (Hunt et al. 2004), which agrees with the repression effect of Hha on the genes related to biofilm. So, it is possible that the nutrient starvation induces the increased expression level of *hha* in late stationary phase. In addition, Hha induces the excision of prophage by regulating excision genes, such as *alpA* at CP4-57 (Kirby et al. 1994) and this excision of prophage can induce cell lysis (Wang et al. 2009a). This may also explain the increased expression level of *hha* over time. Previous microarray data indicated that the transcript abundance of *hha* was

relatively high compared to the average of total gene transcript abundances from exponential phase to 48 h of incubation (Figure 7, 8, 9, 10, Appendix A).

1.4.2 *tomB*

tomB is highly expressed in biofilms compared with planktonic cells (Ren et al. 2004). TomB is an antitoxin to mediate the toxicity of Hha by repressing the expression of *hha* (Garcia-Contreras et al. 2008). This toxin-antitoxin module is highly expressed in stationary phase and in biofilms associated with persister cells (Shah et al. 2006). Hha-TomB could promote conjugation in the presence of conjugative plasmids in cells and further increase biofilm formation by improving cell aggregation and decreasing motility. Under that condition, TomB plays a more important role in the regulation of motility via conjugation (Barrios et al. 2006). The regulation of cell lysis in biofilm by Hha-TomB is also important for biofilm physiology. This toxin-antitoxin module can trigger cellular death in response to starvation. Because the toxin half-life is longer than the antitoxin half-life, after protein synthesis is diminished, the toxin is still free to cause cell death (Garcia-Contreras et al. 2008). This may explain the increased expression level of *tomB* in late stationary phase to adjust cell death in response to nutrient starvation. Microarray data indicated that the transcript abundance of *tomB* was lower than the average of total gene transcript abundances in exponential phase and early stationary phase. However, at 24 h and 48 h of incubation, the transcript abundance of *tomB* increased, exceeding the average level of total transcript abundance (Figure 7, 8, 9, 10, Appendix A).

1.4.3 *emrK*

Multidrug efflux pump genes in *E. coli* can be divided into 5 categories: the ATP binding cassette (ABC) superfamily, major facilitator superfamily (MFS), small MDR (SMR) family, resistance-nodulation-cell division (RND) family and multidrug and toxic compound extrusion (MATE) family (Lubelski et al. 2007). They can be also divided into two classes: proton motive force (PMF) pumps or sodium ion motive force (MATE family) pumps. EmrK belongs to the MFS family and is a kind of proton motive force (PMF) pumps (Matsumura et al. 2011). The expression of *emrK* increases during biofilm formation. It is possible that efflux pumps export or import some substrate important for biofilm formation (Soto 2013). The expression of *emrK* is induced in the presence of H_2O_2 (Han et al. 2010), which is produced when cells grow under aerobic conditions. Indole is metabolically synthesized in *E. coli* and is regarded as a stationary phase extracellular signal for many bacteria (Snell 1975). The synthesis of indole also induces the expression of *emrK* in *E. coli* (Hirakawa et al. 2005). Based on the previous microarray data, the transcript abundance of *emrK* was kept at a low level from exponential phase to 24 h of incubation. At 48 h of incubation, the transcript abundance exceeded the average level of total gene transcript abundance (Figure 7, 8, 9, 10, Appendix A).

1.4.4 *yafY*

YafY is an inner membrane lipoprotein and can strongly induce *degP* expression by overproduction. DegP is a periplasmic protease that digests unfolded proteins and

important for bacteria growing at high temperatures (Strauch et al. 1989). Although the precise function of YafY is still not known, it is possible that YafY monitors certain cell envelop stress stimuli (Miyadai et al. 2004). That may explain the increased expression level of *yafY* over time. In addition, the expression of *yafY* is regulated by RpoE (Onufryk et al. 2005). According to the previous microarray result, the transcript abundance of *yafY* was lower than average of total gene transcript abundances from exponential phase to 24 h of incubation and increased to the average of total gene transcript abundance from 24h to 48h of incubation (Figure 7, 8, 9, 10, Appendix A).

1.4.5 *ytfE*

YtfE is involved in the repair of degraded iron-sulfur clusters in *E. coli*. Protein with iron-sulfur clusters is one of the most ubiquitous metalloproteins in nature and plays an important role in cell metabolism (Todorovic et al. 2008). YtfE can enhance bacteria resist nitrosative and oxidative stress (Justino et al. 2007). Because the repair of nitrosative and oxidative stress requires reinsertion of iron, it is possible that YtfE helps these metalloproteins recruit and integrate ferrous iron (Justino et al. 2007). In addition, the expression of *ytfE* is also induced by nitric oxide (NO) because NO damages the iron-sulfur proteins (Justino et al. 2005). Bacteria undergo oxidative stress in stationary phase due to the accumulation of ROS, such as H_2O_2 and reactive nitric oxide (Apel and Hirt 2004). This may be the reason for the increased expression level of *ytfE* between 24h and 48h of incubation. Previous microarray data indicated that the transcript abundance of

ytfE was higher than the average of total gene transcript abundances at different time points and kept increasing over time (Figure 7, 8, 9, 10, Appendix A).

1.4.6 *mcrC*

McrC is a subunit of McrBC, which is a restriction enzyme that can specifically recognize and cleave hemimethylated or fully methylated DNA with the recognition sequence R^mCN_{-30}~~2000R^mC. This cleavage is strictly dependent on GTP hydrolysis (Ross et al. 1989). McrC contains a catalytic center for DNA cleavage. A multiple sequence alignment shows that a sequence motif, the PD…D/EXK is conserved in McrC in many species (Pieper and Pingoud 2002). The methylation of DNA influences chromosome replication, mismatch repair (Wion and Casadesus 2006) as well as some gene expression (Falker et al. 2006). This may explain the increased expression of *mcrC* over time. Previous microarray data indicated the transcript abundance of *mcrC* was lower than the average of total gene transcript abundances from exponential phase to 24 h of incubation. However, at 48 h of incubation, the transcript abundance of *mcrC* increased to exceed the average of total transcript abundance (Figure 7, 8, 9, 10, Appendix A).

1.4.7 *yfdK*

yfdK is located at prophage CPS-53 and enhances the resistance to oxidative stress (Wang et al. 2010). However, the precise function of YfdK is still unknown. In order to resist the oxidative stress in late stationary phase, the expression of *yfdK* may increase over time. Previous microarray data indicated that the transcript abundance of *yfdK*

increased over time and exceeded the average of total gene transcript abundance at 48 h of incubation (Figure 7, 8, 9, 10, Appendix A).

1.4.8 *yhcE*

yhcE is homologous to the carboxy-terminal domain of the homocysteine methyltransferases (*metE* family). Contrary to the classical methionine synthase MetE, YhcE converts methionine to cysteine in *L. lactis* (Sperandio et al. 2005). The expression of *yhcE* increases in the absence of cysteine and under high content of sulfide in medium (Sperandio et al. 2005). YhcE is involved in the sulfur metabolism and the production of cysteine can reduce ferric iron, which permits free iron to redox cycle in the presence of H_2O_2 (Park and Imlay 2003). The oxidative stress in late stationary phase may induce the expression of *yhcE*. Based on the microarray data, the transcript abundance of *yhcE* was lower than the average of total gene transcript abundances from exponential phase to 24 h of incubation and increased to exceed the average of total gene transcript abundance at 48 h of incubation (Figure 7, 8, 9, 10, Appendix A).

1.4.9 *bolA*

BolA regulates the expression of genes involved in the morphology of the cells in stationary phase and stress conditions (Aldea et al. 1988). The overexpression of *bolA* can regulate elongation, cell septation (Aldea et al. 1988) and the formation of biofilm (Vieira et al. 2004). BolA produces round-shaped cells by mediating the expression of *dacA*, encoding PBP5, *dacC*, encoding PBP6 and *ampC*, which are all involved in the murein

metabolism (Santos et al. 2002). BolA also changes the properties of cell outer membrane (Santos et al. 2002). The overexpression of *bolA* reinforces integrity of outer membrane by increasing the resistance to surface-active agents, such as SDS, and high molecular-mass antibiotics, such as vancomycin. In addition, the overexpression of *bolA* can increase the ratio of OmpC/OmpF to decrease the cell membrane permeability. All these changes protect cells against stress conditions (Santos et al. 2002). The expression of *bolA* is also regulated by RpoS (Lange and Hengge-Aronis 1991). The transcript abundance of *bolA* was lower than the average of total gene transcript abundances in exponential phase and 48 h of incubation but became high in stationary phase and 24 h of incubation (Figure 7, 8, 9, 10, Appendix A).

1.4.10 *osmY*

OsmY is a periplasmic protein and is induced under hyperosmotic stress and stationary phase (Yim and Villarejo 1992). The expression of *osmY* is positively regulated by RpoS and repressed by Lrp, cAMP-CRP and IHF (Lange et al. 1993). RpoD also regulates the expression of *osmY* (Lange et al. 1993). However, under hyperosmotic stress, such as in high concentration of potassium glutamate, the transcription is mainly regulated by RpoS (Ding et al. 1995). The transcript abundance of *osmY* was much higher than the average of total gene transcript abundances from exponential phase to 48 h of incubation compared with other genes (Figure 7, 8, 9, 10, Appendix A).

Quantitative real-time PCR (qPCR) was used to quantify gene expression in late stationary phase. Compared to microarray, qPCR is more sensitive in producing reliable

and rapid quantification results. There are two common methods to measure gene expression by qPCR: absolute and relative quantification. Absolute qPCR measures the exact copy number of target genes by the construction of a standard curve (Fu et al. 2009). The relative quantification method can be further categorized into two types: the ΔΔCt method and relative standard curve method. In relative quantification, reference genes are introduced to normalize the expression of target genes. *rrsA*, encoding 16S ribosomal RNA, is widely used as a reference gene in qPCR. The expression of *rrsA* is constant across different conditions (Zhou et al. 2011). In this study, the relative standard method was used to quantify target gene expression at different time points. In addition, RNA-seq, a comprehensive and accurate transcriptome analysis tool (Giannoukos et al. 2012), was also used to identify gene expression in late stationary phase. RNA-seq has the advantages of high sensitivity to identify and quantify low abundance transcripts, and low requirement for input RNA (Wang et al. 2009b). Dependent on different purposes, a proper sequencing depth is chosen. For instance, in order to create a more comprehensive transcriptome profile, the number of reads per library and the reads representing low abundance transcripts should be increased, which can be achieved by depleting ribosomal RNA (rRNA), because rRNA accounts for 80-90% bacterial transcriptome. To compare the expression of same gene under different conditions, the breadth (the number of sequenced samples) is more important. It is more valuable to provide additional biological replicates than to detect low abundance transcripts (Haas et al. 2012).

In this study, qPCR and RNA-seq was performed to identify the expression of genes in *E. coli* in late stationary phase. The expression of *hha, tomB, emrk, bolA* and

osmY in exponential phase, early stationary phase, and at 24 h and 48 h of incubation was quantified by qPCR to validate previous microarray data in the lab. RNA-seq was also performed to profile global gene expression, especially rRNA and genes with abundant transcripts at 24 h and 48 h of incubation.

Chapter 2 Methodology

2.1 Strains, media and growth conditions

E. coli MG1655 strain was used in this study. Stocked *E. coli* MG1655 was streaked on an LB agar plate and incubated at 37˚C overnight. Independent colonies were inoculated in 50 ml of fresh LB and grown aerobically at 37℃ with shaking at 200 rpm. Overnight cultures were diluted into fresh media at a starting OD_{600} of 0.0001 to allow cells to grow at least 10 generations prior to RNA isolation in exponential phase. Cultures grown in 5 biological replicates were sampled at OD_{600} of 0.3 (exponential phase) and OD_{600} of 1.5 (early stationary phase), 24 h and 48 h of incubation. Growth was monitored spectrophotometrically at OD_{600}.

2.2 RNA preparation

RNA was isolated using hot phenol-chloroform method (Dong et al. 2008). Bacterial cultures were treated with half volume of boiling lysis solution (2% SDS and 16 mM EDTA) and same volume of pre-heated acidic phenol. After isolation by using three

hot phenol-chloroform extraction steps, DNase I was added into RNA samples. Hot phenol-chloroform extraction was done three more times to remove DNase.

To reduce RNA loss from the extraction steps, the hot phenol-chloroform method was modified by adding DNase after the first extraction. 80U of DNase I was added to each sample and incubated at 37℃ for 20 min. Then, three extraction steps with hot phenol-chloroform were performed. RNA was also isolated with Total RNA Purification Kit (Norgen Biotek, #17200). During extraction, DNase I was added into samples on column according to the instructions from Norgen Biotek. RNA concentration was measured with the Qubit fluorometer (Invitrogen). The integrity of RNA was checked on 0.8% agarose gel. RNA samples were stored at -80˚C. RNA recovery was calculated by dividing the recovered RNA amount with the yield RNA amount.

2.3 DNase titration

To measure the efficiency of DNase on the removal of genomic DNA, 0 U, 10 U, 20 U, 40 U, 60 U and 80 U of DNase I was treated with RNA samples on the column when RNA was extracted from cells of OD_{600} of 1.0 at 24 h of incubation according to the instruction from Norgen Biotek. DNA and RNA concentration was measured with the Qubit fluorometer.

2.4 Lysis solution effect on the measurement of yield RNA concentration

To better estimate the recovery of extracted RNA, after cells were lysed, the concentration of PLYR (Post lysis yield RNA) was measured before purification (Figure

1). To examine the effect of lysis solution on RNA concentration measurement, 800 ng/ml of RNA was resuspended in 20%, 40%, 60%, 80% and 100% lysis solution and the concentration was measured with the Qubit fluorometer. 800 ng/ml of RNA in DEPC water was used as a positive control. The standard curve of RNA concentration measurement as a function of the percentage of lysis solution was then generated.

2.5 DNA preparation

E. coli MG1655 strain was streaked out on an LB agar plate. A single colony was inoculated into 10 mL of LB and grown aerobically at 37℃ with shaking at 200 rpm overnight. DNA extraction was performed using Bacterial Genomic DNA Isolation Kit (Norgen Biotek, #17900). DNA concentration was measured with the Qubit fluorometer.

2.6 cDNA synthesis

First strand cDNA was synthesized using cDNA Synthesis Kit (Bio-Rad). Reverse transcription mixtures (20 μl) containing 4 μl of RT supermix, 160 ng of RNA and H_2O were used. The final concentration of cDNA was 8 ng/μl. cDNA was stored at -20℃.

2.7 Quantitative real time PCR (qPCR) assay

Primers amplifying 100-300 bp of target genes were designed by Primer 3 or obtained from Sarah Chiang. qPCR mixtures (10 μl) containing 8 ng of cDNA or RNA (NRT), 4 μl of 2X SsoFast^{TM} EvaGreen Supermix (Bio-Rad), 0.2 μM of forward and reverse primer and 4 μl of H_2O were used.

The PCR cycling conditions comprised of enzyme activation at 95 ℃ for 2 min, followed by 40 cycles of denaturation at 95 ℃ for 5 s and annealing/extension at 62 ℃ for 3 s. A melt curve assay was then performed with increments of 0.3 ℃ ranged from 55-95 ℃ for 5 s. qPCR was performed on the CFX96 Touch™ Real-Time PCR Reaction System (Bio-Rad).

2.8 Generation of standard curves

To determine qPCR amplification efficiency (E) and the dynamic range, the standard curve of each gene was produced. Serial dilutions of genomic DNA were used as the templates. The standard curve was generated by plotting Cq (quantification cycle) values against the log values of the copy numbers of tested genes.

2.9 Data analysis by the relative standard curve method

The copy numbers of the target and reference genes were obtained by the plotting the Cq values against the standard curve. *rrsA* was used as a reference gene to normalize the difference of input RNA quantity among samples (Kobayashi et al. 2006). To normalize target gene expression, the average copy numbers of target genes were divided with the average copy numbers of *rrsA* at each time point. cDNA of exponential phase was used as a calibrator sample. cDNA of early stationary phase, 24 h and 48 h of incubation was used as the test samples. To calculate the fold changes between growth phases, the normalized copy numbers of the target genes in tested samples were divided

with that in calibrator samples. Standard error (S.E.) was calculated. Data was analyzed on the Bio-Rad CFX Manager™.

2.10 RNA-Seq

1 μg of RNA extracted from cells of OD_{600} of 1.0 at 24 h and 48 h of incubation was sent to Farncombe Metagenomics Facility (McMaster University, Hamilton). Strand-specific cDNA library was prepared without rRNA depletion.

Table 1. Oligonucleotides used in this study.

Gene	Primer sequence (5'-3')	Amplicon size (bp)
rrsA	F: GTA ATC GTG GAT CAG AAT GC	160
	R: ACC TTG TTA CGA CTT CAC C	
hha	F: CTG GAG CGT GTT ATC GAG	114
	R: GTC GTA CAG TTT ATT CAT GGT C	
tomB	F: TAA GCT CAT TGA GCA GAT CG	119
	R: AAC GGA ATA GTC GAT TAC CTG	
emrK	F: ATT GAA TGC CGC TAT CCA	183
	R: CCG ACC TGA ACA CTT CTC	
osmY	F: ATATCGTCCCTTCCCGTC	127
	R: TTCACACCATCTACCGCT	
bolA	F: CGTATTCCTCGAAGTAGTGG	158
	R: TAGAGAGTTCCTCCGCTAAA	

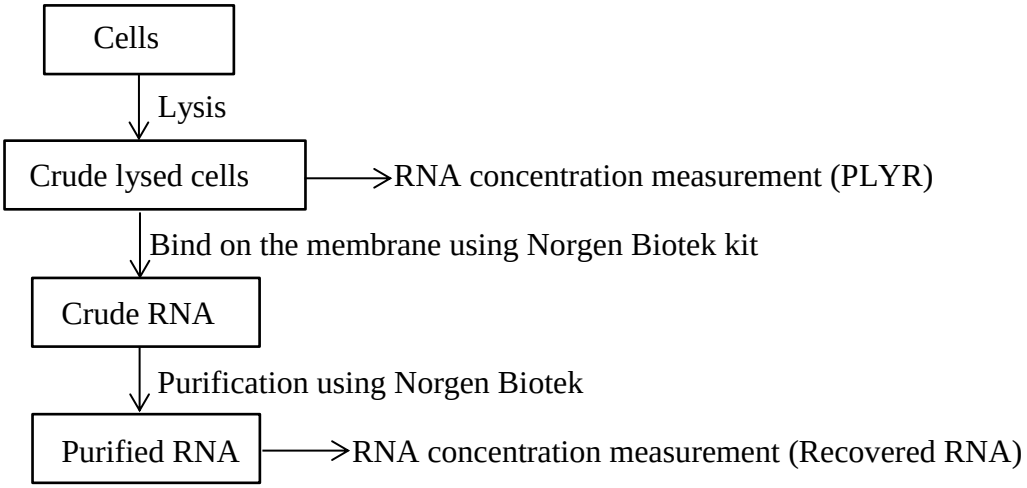


Figure 1. Flowchart for determination of RNA recovery when RNA was extracted using a total RNA isolation kit (Norgen Biotek, #17200). PLYR (post lysis yield RNA) is the crude RNA after cell lysis.

Chapter 3. Results

3.1 Cell growth

To quantify target gene expression at different time points, samples were collected from bacterial cultures from exponential phase to 48 h of incubation. Overnight *E. coli* MG1655 cultures were subcultured into fresh media at a starting OD_{600} of 0.0001 to fully dilute existing RNA and allow cells to grow at least 10 generations prior to exponential phase. Cultures were grown at 37˚C with shaking at 200 rpm and the growth was monitored spectrophotometrically at $OD_{600.}$ At 24 h of incubation, OD_{600} was around 4.66 and it dropped down to 3.78 at 48 h of incubation (Table 2). *E. coli* cells at 48 h of incubation were plated on an LB agar plate to monitor contamination.

3.2 RNA extraction from *E. coli* in late stationary phase using hot phenol-chloroform method

To extract RNA from *E. coli* in late stationary phase, a large-scale hot phenol-chloroform method was first performed. In the lab, RNA has been successfully extracted by this method from cultures in exponential phase and early stationary phase (Patten et al. 2004). 24 h and 48 h bacterial cultures (4 ml) were immediately mixed with 2 ml of boiling 2% SDS/ 16Mm EDTA and equal volume (6 ml) of hot phenol (pH 4.3, 65˚C). RNA concentration was measured with the Qubit fluorometer and the yield RNA amount was calculated by multiplying 10 pg (RNA amount per cell) with cell numbers (OD_{600} of 1.0 = 8 x 10^8 cell/ml) (Sasagawa et al. 2013). RNA recovery was calculated by dividing

the recovered RNA amount with the yield RNA amount. The recovered RNA at 24 h of incubation was 52 μg, corresponding to 5% recovery. At the 48h of incubation, the recovered amount was around 56 μg, corresponding to 4% recovery. During extraction, RNA was easily degraded. Therefore, a small-scale hot phenol-chloroform method was performed to extract RNA from 24 h and 48 h bacterial cultures using microcentrifuge tubes. 1.5 ml of bacterial cells was treated with 500 μl of boiling lysis solution (2% SDS and 16 mM EDTA) and 1 ml of pre-heated acidic phenol to inactivate RNase and repress RNA degradation. After electrophoresis, the genomic DNA band could be visualized on a gel. DNase was added to remove genomic DNA. Hot phenol-chloroform extractions were performed three times to remove DNase. After that, the RNA concentration was too low to be detected with the Qubit fluorometer.

To avoid extra extraction steps to remove DNase, the hot phenol-chloroform method was modified by directly adding 80 U of DNase after the first extraction and incubating the samples at 37˚C for 20 min. The recovered RNA amount from 24 h bacterial cultures was around 20 μg, corresponding to 3.65% RNA recovery. The recovered RNA amount from 48 h bacterial cultures was around 10 μg, corresponding to 2.50% RNA recovery (Table 3).

3.3 DNase titration

To solve the low RNA recovery in late stationary phase and genomic DNA contamination issue, RNA was then extracted with a commercial kit (Norgen Biotek). 80 U of DNase I was required according to the instructions from Norgen Biotek. To measure

the efficiency of DNase, DNase titration was performed prior to RNA extraction from cells from exponential phase to 48 h of incubation. 24 h bacterial culture was used in this experiment. RNA was extracted from cells of OD_{600} of 1.0. Different amounts of DNase I was added into samples on the column according to the instructions. 20 U of DNase I was sufficient to remove genomic DNA, with the final DNA/RNA concentration ratio at 3% (Table 4). 97% genomic DNA was removed. DNase treatment did not affect the recovery of RNA (Table 4). 22% RNA was recovered when RNA was extracted from cells at 24 h of incubation.

3.4 RNA extraction with the commercial kit (Norgen Biotek, #17200)

Subsequently, RNA was extracted from cells of OD_{600} of 1.0 in exponential phase, early stationary phase, 24 h and 48 h of incubation using the RNA isolation kit. After isolation, the integrity of RNA was checked on 0.8% agarose gel and the quantity of RNA was measured with the Qubit fluorometer. The electrophoresis results indicated the integrity of RNA (Figure 1). DNA concentration was also measured with the Qubit fluorometer. Genomic DNA contamination ratio was calculated by dividing DNA concentration with RNA concentration. 80 μg of RNA was used as the yield RNA amount when RNA was extracted from cells of OD_{600} of 1.0, which was calculated by multiplying 10 pg (RNA amount per cell) with cell numbers (8×10^8 cells) (Sasagawa et al. 2013). Standard deviation was calculated to represent the variability of each extraction. The RNA recovery in exponential phase, early stationary phase, 24 h and 48 h of incubation was around 58.3%, 48.5%, 27.3% and 1.6% respectively. The genomic DNA

contamination ratio in exponential phase, early stationary phase and 24 h of incubation was around 2.0%. In 48 h RNA samples, the ratio increased to 9.2% (Table 5).

3.5 Lysis solution's effect on the measurement of RNA concentration

To better calculate RNA recovery, the PLYR (Post lysis yield RNA) amount was examined by measuring RNA concentration after cell lysis. The PLYR amount was even lower than the recovered RNA amount based on the RNA concentration measurement from the Qubit fluorometer (Table 6). To test whether lysis solution affected the measurement and how much the effect was, RNA with the same concentration was resuspended in different percentages of lysis solution. RNA concentration was measured with the Qubit fluorometer. RNA in DEPC water was used as a positive control. The measurement result of RNA concentration decreased sharply with the increased percentage of lysis solution. The concentration of RNA in 80% and 100% lysis solution could not be detected due to the limitation that the Qubit fluorometer could not detect RNA at concentrations lower less than 20 ng/ml. The concentration of RNA in 20%, 40% and 60% lysis solution was 207 ng/ml, 47 ng/ml and 21 ng/ml (Figure 2).

3.6 Amplification efficiency and specificity tests of target and reference genes

Before quantification analysis, the efficiency of PCR should be determined, because it indicates the problems with the qPCR reaction. 90-110% efficiencies are required for qPCR assay in microarray validation (Morey et al. 2006). The amplification efficiency can be determined by constructing a standard curve. 5-fold serial dilutions of

genomic DNA starting from 43 ng were used as the templates. The standard curve can also determine the linear dynamic range. In this study, the linear dynamic ranges were between 10^2 to 10^7 gene copies. The amplification efficiency of *rrsA, hha, tomB, emrK, bolA* and *osmY* was 93.6%, 106.7%, 97%, 96.1%, 91.3% and 101.5%, respectively (Table 7, Figure 1, 2, 3, 4, 5, 6, Appendix A) . The coefficient of determination (R^2) of *rrsA, hha, tomB, emrK, bolA* and *osmY* was 0.989, 0.985, 0.994, 0.985, 0.987 and 0.980 (Table 7, Figure 1, 2, 3, 4, 5, 6, Appendix A). The construction of a melt curve from 55-95 ℃ with increments of 0.3 ℃ identified the amplified target products (Figure 17, 18, 19, 20, 21 and 22, Appendix A).

3.7 Technical replication test

Due to the limited size of the plate, there was not enough space to run technical replicates of each biological replicate at one time. The technical replication test was firstly performed. The expression of *tomB* in exponential phase, early stationary phase, 24 h and 48 h of incubation was assayed from one biological replicate using the relative standard curve method in three technical replicates. The standard deviations were calculated to represent the variability of each technical operation. qPCR results indicated the expression of *tomB* decreased from exponential phase to early stationary phase and increased from early stationary phase to 48 h of incubation. From 24 h to 48 h of incubation, the expression level increased 10-fold (Table 8).

3.8 Quantification of target gene expression at different time points

To identify the preferentially expressed genes in *E. coli* in late stationary phase, the expression of *hha, tomB, emrK, osmY* and *bolA* at different time points was measured by the relative standard method of qPCR. Three biological replicates were assayed and the standard errors (S.E) were calculated to represent the deviation of the sampling distribution. qPCR results indicated that the expression of *hha* maintained similar levels from exponential phase to early stationary phase and then increased from early stationary phase to 48 h of incubation, especially from early stationary phase to 24 h of incubation. During that period of time, the expression of *hha* increased 108-fold (Table 9). qPCR results indicated that from exponential phase to early stationary phase, the expression of *tomB* increased 2-fold. From early stationary phase to 24 h of incubation, *tomB* expression level increased almost 60-fold. From 24 h to 48 h of incubation, the expression level of *tomB* increased around 4-fold (Table 10). As for the expression of *emrK*, qPCR results indicated that from exponential phase to early stationary phase, the expression of *emrK* increased 3-fold. From early stationary phase to 24 h of incubation, it increased 11-fold. From 24 h to 48 h of incubation, the expression level increased 7-fold (Table 11). qPCR results indicated that the expression of *osmY* and *bolA* increased from exponential phase to early stationary phase and then decreased 17-fold and 2-fold respectively from early stationary phase to 24 h of incubation. From 24 h to 48 h of incubation, the expression of *osmY* and *bolA* increased 3-fold (Table 12, 13). The Cq values of each gene were within the dynamic range of the standard curve, showing the validity of the data (Figure 11, 12, 13, 14, 15 and 16, Appendix A) (Bustin et al. 2009).

Table 2. Cell growth from exponential phase to 48 h of incubation.

Overnight *E. coli* MG1655 cultures were subcultured into fresh media at a starting OD_{600} of 0.0001. Cell growth was monitored spectrophotometrically at OD_{600}

Time (h)	Growth phase	OD_{600}
5	Exponential	0.32 ± 0.02
7	Early stationary	1.52 ± 0.04
24	1 day	4.66 ± 0.12
48	2 days	3.78 ± 0.06

Table 3. RNA isolation by a modified small-scale hot phenol-chloroform method.

RNA was extracted from *E. coli* MG1655 at 24 h and 48 h of incubation. RNA was isolated by a modified small-scale hot phenol-chloroform method. 80 U of DNase I was added after first extraction to avoid extra extraction steps to remove DNase (i.e. steps used in the traditional hot phenol-chloroform method).

Time (h)	Growth phase	Recovered RNA concentration (μg/ul)	Recovered RNA amount (μg)	Yield RNA [b] amount (μg)	RNA recovery [a] (%)
24	1 day	0.99 ± 0.13	19.82 ± 2.58	540	3.65 ± 0.45
48	2 days	0.53 ± 0.11	10.62 ± 2.18	432	2.50 ± 0.50

[a] RNA recovery was determined by using recovered RNA amount divided with yield RNA amount.

[b] Yield RNA amount was determined by multiplying10 pg(RNA amount per cell) with cell numbers.

Table 4. DNase effect on genomic DNA removal and RNA recovery.

RNA was extracted from cells at OD_{600} of 1.0 at 24 h of incubation. Different amounts of DNase I were treated with RNA samples on column according to the instruction from Norgen Biotek.

DNase amount (U)	RNA concentration (ng/μl)	RNA recovery (%) [a]	DNA concentration (ng/μl)	Genomic DNA contamination (%) [b]
0	460	22	78	17
10	540	22	24	4
20	576	23	14	3
40	448	18	12	3
60	456	18	12	3
80	373	14	10	3

[a] RNA recovery was determined by using recovered RNA amount divided with yield RNA amount (80 μg).

[b] Genomic DNA contamination was determined by dividing DNA concentration with RNA concentration.

Table 5. RNA isolation with the commercial kit (Norgen Biotek, #17200).

RNA was extracted from *E. coli* MG1655 in exponential phase, early stationary phase, 24 h and 48 h of incubation. A commercial RNA isolation kit (Norgen Biotek, #17200) was used to isolate RNA. During isolation, on-column DNase treatment was performed.

Time (h)	Growth phase	RNA concentration (ng/μl)	RNA amount (ug)	Recovered RNA (%) [a]	Genomic DNA contamination (%) [b]
5	Exponential	1203.6 ± 56.5	47.8 ± 2.3	58.3 ± 4.1	2.0 ± 0.0
7	Early stationary	932.4 ± 22.1	37.0 ± 1.1	48.5 ± 1.9	2.0 ± 0.0
24	1 day	30.6 ± 11.5	12.2 ± 0.5	27.3 ± 2.0	2.4 ± 0.4
48	2 days	30.8 ± 1.4	1.0 ± 0.0	1.6 ± 0.1	9.2 ± 0.4

[a] RNA recovery was determined by using recovered RNA amount divided with yield RNA amount (80 μg) in each sample.

[b] Genomic DNA contamination was determined by using DNA concentration divided with RNA concentration in each sample.

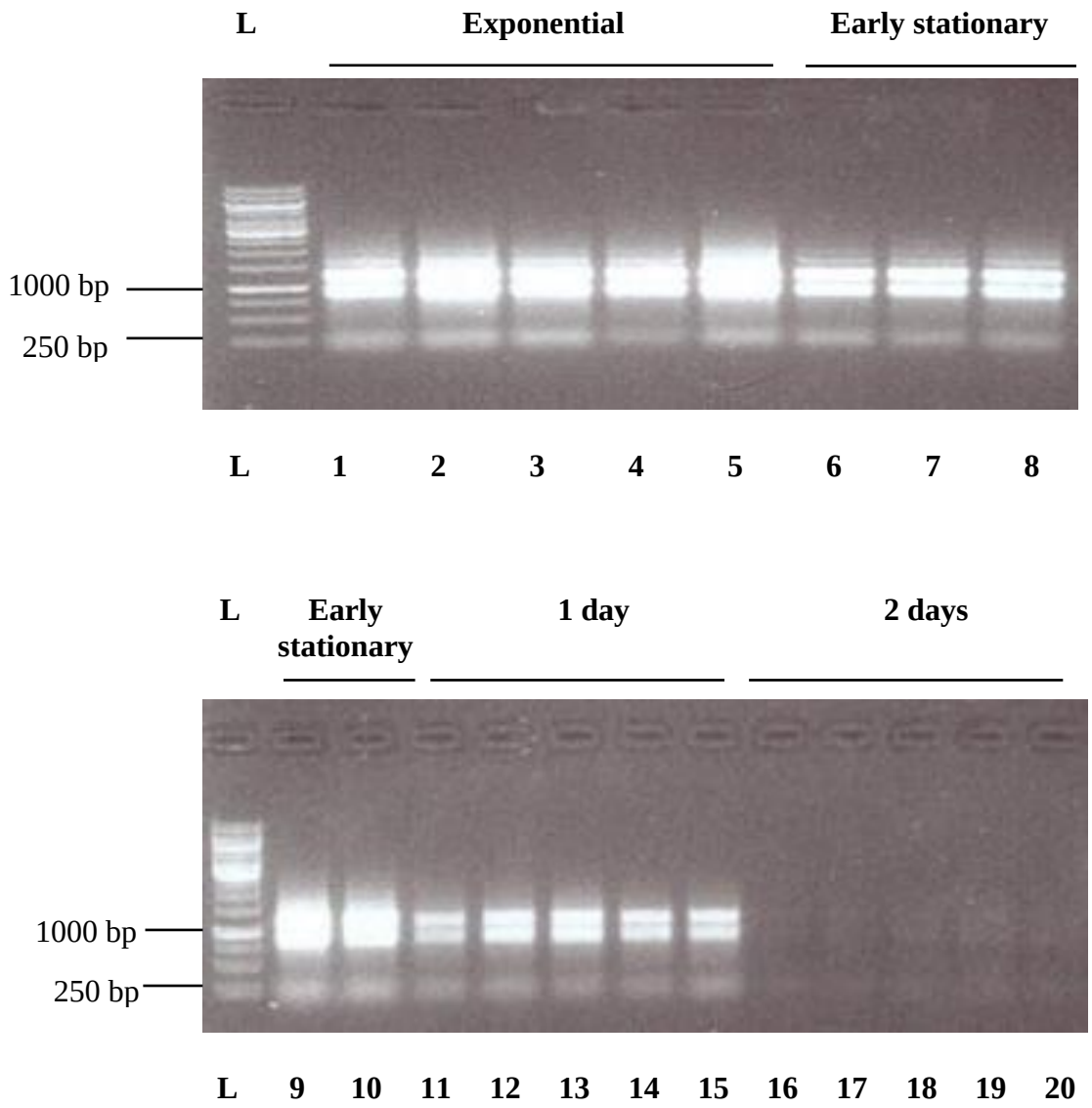


Figure 2. RNA integrity check on 0.8% agarose gel.

RNA was extracted from *E. coli* MG1655 from exponential phase to 48 h of incubation in five biological replicates using a commercial kit (Norgen Biotek, #17200). Lane 1-5: RNA extracted from cells in exponential phase; Lane 6-10: RNA extracted from cells in early stationary phase; Lane 11-15: RNA extracted from cells at 24 h of incubation; Lane 16-20: RNA extracted from cells at 48 h of incubation. L is a 1 kb ladder.

Table 6. Comparison between PLYR (post lysis yield RNA) and recovered RNA.

The PLYR (post lysis yield RNA) and recovered RNA from exponential phase and 48 h bacterial cultures were compared to examine RNA loss during RNA isolation by the commercial kit (Norgen Biotek, #17200).

Time (h)	Growth phase	PLYR [a] concentration (ng/ul)	PLYR [a] amount (ng)	Recovered RNA [b] concentration (ng/ul)	Recovered RNA [b] amount (ng)
5	Exponential	6.62	2648	1280	51200
48	2 days	0.33	132	37	1488

[a] PLYR (post lysis yield RNA) : RNA in the cells after lysis.

[b] Recovered RNA : extracted RNA after isolation.

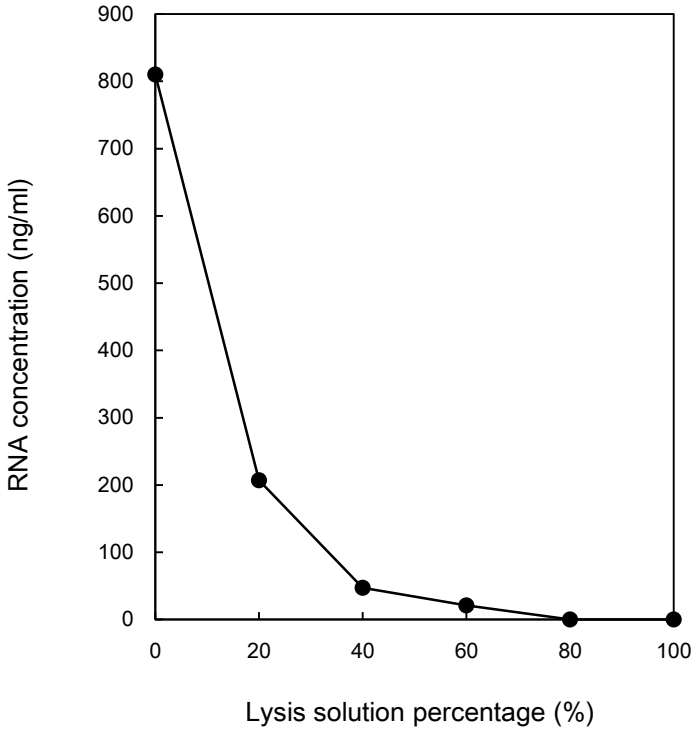


Figure 3. Lysis solution effect on the measurement of RNA concentration.

RNA with the same concentration (800 ng/ml) was resuspended in 20%, 40%, 60%, 80% and 100% lysis solution. RNA concentration was then measured with the Qubit fluorometer to examine the effect of lysis solution on RNA concentration measurement results.

Table 7. Standard curves of target and reference genes.

The amplification efficiencies of target and reference gene were tested by the construction of standard curves. The coefficient of determination was exported from Bio-Rad CFX Manager™.

Target/reference gene	Slope of the standard curve [a]	E [b]	R^2 [c]
rrsA	-3.485	93.6%	0.989
hha	-3.299	106.7%	0.985
tomB	-3.395	97.0%	0.994
emrK	-3.419	96.1%	0.985
bolA	-3.550	91.3%	0.987
osmY	-3.222	104.5%	0.980

[a] Standard curves were produced by plotting Cq values (quantification cycle values) against log values of the gene copy numbers.

[b] PCR amplification efficiency was calculated by using $10^{(-1/\text{slope of the standard curve})}$ -1 from the standard curve.

[c] Coefficients of determination

Table 8. Technical replication test.

The technical replication test was performed to quantity *tomB* expression in exponential phase, early stationary phase, 24 h and 48 h of incubation. qPCR was assayed from one biological replicate in three technical replicates by the relative standard curve method. The expression of *tomB* was normalized by *rrsA*. The expression level change between growth phases was represented by fold change.

Time (h)	Growth phase	Expression (log_2 transformed)		Normalized expression value [a]	Fold change [b]	
		tomB	*rrsA*		qPCR	Microarray
5	Exponential	14.8 ±0.3	21.9 ±0.4	0.007	1.0	1.0
7	Early stationary	13.6 ±0.5	21.3 ±0.7	0.005	0.7	1.0
24	1 day	11.4 ±0.2	17.7 ±0.7	0.006	0.8	4.5
48	2 days	16.0 ±0.8	19.6 ±1.2	0.076	10.5	24.0

[a] The normalized expression of *tomB* in each sample was calculated using the expression of *tomB* divided by the expression of *rrsA*.

[b] Fold change was determined by using the normalized expression of *tomB* at each time point divided by the normalized expression of *tomB* in exponential phase.

Table 9. Quantification of *hha* expression from exponential phase to 48 h of incubation.

The expression of *hha* in exponential phase, early stationary phase, 24 h and 48 h of incubation was quantified by qPCR relative standard curve method. The expression of *hha* was normalized by *rrsA*. The expression level change between growth phases was represented by fold change.

Time (h)	Growth phase	Expression (log_2 transformed)		Normalized expression value [a]	Fold change [b]	
		bolA	*rrsA*		qPCR	Microarray
5	Exponential	14.6 ± 0.2	22.1 ± 0.4	0.005	1	1
7	Early stationary	18.2 ± 0.4	21.9 ± 0.1	0.083	16	20
24	1 day	15.9 ± 0.3	20.7 ± 0.2	0.037	7	345
48	2 day	15.2 ± 0.3	18.6 ± 0.4	0.097	19	3

[a] The normalized expression of *hha* in each sample was calculated using the expression of *hha* divided by the expression of *rrsA*.

[b] Fold change was determined by using the normalized expression of *hha* at each time point divided by the normalized expression of *hha* in exponential phase.

Table 10. Quantification of *tomB* expression from exponential phase to 48 h of incubation.

The expression of *tomB* in exponential phase, early stationary phase, 24 h and 48 h of incubation was quantified by qPCR relative standard curve method. The expression of *tomB* was normalized by *rrsA*. The expression level change between growth phases was represented by fold change.

Time (h)	Growth phase	Expression (log_2 transformed)		Normalized expression value [a]	Fold change [b]	
		tomB	*rrsA*		qPCR	Microarray
5	Exponential	14.9 ± 0.9	22.5 ± 0.3	0.006	1	1
7	Early stationary	16.7 ± 0.1	22.7 ± 0.3	0.015	2	1
24	1 day	19.5 ± 1.3	21.1 ± 0.1	0.725	118	5
48	2 days	20.7 ± 0.5	19.1 ± 0.3	3.180	519	24

[a] The normalized expression of *tomB* in each sample was calculated using the expression of *tomB* divided by the expression of *rrsA*.

[b] Fold change was determined by using the normalized expression of *tomB* at each time point divided by the normalized expression of *tomB* in exponential phase.

Table 11. Quantification of *emrK* expression from exponential phase to 48h of incubation.

The expression of *emrK* in exponential phase, early stationary phase, 24 h and 48 h of incubation was quantified by qPCR relative standard curve method. The expression of *emrK* was normalized by *rrsA*. The expression level change between growth phases was represented by fold change.

Time (h)	Growth phase	Expression (log_2 transformed)		Normalized expression value [a]	Fold change [b]	
		emrK	*rrsA*		qPCR	Microarray
5	Exponential	8.2 ± 0.3	19.8 ± 0.4	0.0003	1	1.0
7	Early stationary	9.8 ± 0.5	20.4 ± 0.2	0.0007	3	1.4
24	1 day	10.7 ± 1.1	18.3 ± 0.4	0.0088	34	2.7
48	2 day	12.4 ± 0.5	16.6 ± 0.4	0.0562	216	13.1

[a] The normalized expression of *emrK* in each sample was calculated using the expression of *emrK* divided by the expression of *rrsA*.

[b] Fold change was determined by using the normalized expression of *emrK* at each time point divided by the normalized expression of *emrK* in exponential phase.

Table 12. Quantification of *osmY* expression from exponential phase to 48 h of incubation.

The expression of *osmY* in exponential phase, early stationary phase, 24 h and 48 h of incubation was quantified by qPCR relative standard curve method. The expression of *omsY* was normalized by *rrsA*. The expression level change between growth phases was represented by fold change.

Time (h)	Growth phase	Expression (log_2 transformed)		Normalized expression value [a]	Fold change [b]	
		osmY	*rrsA*		qPCR	Microarray
5	Exponential	13.3 ± 0.8	22.2 ± 0.2	0.003	1	1
7	Early stationary	20.3 ± 0.8	22.2 ± 0.3	0.384	135	14
24	1 day	15.3 ± 0.2	20.8 ± 0.1	0.022	8	52
48	2 day	14.4 ± 0.5	18.4 ± 0.4	0.061	24	8

[a] The normalized expression of *osmY* in each sample was calculated using the expression of *osmY* divided by the expression of *rrsA*.

[b] Fold change was determined by using the normalized expression of *osmY* at each time point divided by the normalized expression of *osmY* in exponential phase.

Table 13. Quantification of *bolA* expression from exponential phase to 48h of incubation.

The expression of *bolA* in exponential phase, early stationary phase, 24 h and 48 h of incubation was quantified by qPCR relative standard curve method. The expression of *bolA* was normalized by *rrsA*. The expression level change between growth phases was represented by fold change.

Time (h)	Growth phase	Expression (log_2 transformed)		Normalized expression value [a]	Fold change [b]	
		bolA	*rrsA*		qPCR	Microarray
5	Exponential	14.6 ± 0.2	22.1 ± 0.4	0.005	1	1
7	Early stationary	18.2 ± 0.4	21.9 ± 0.1	0.083	16	20
24	1 day	15.9 ± 0.3	20.7 ± 0.2	0.037	7	345
48	2 day	15.2 ± 0.3	18.6 ± 0.4	0.097	19	3

[a] The normalized expression of *bolA* in each sample was calculated using the expression of *bolA* divided by the expression of *rrsA*.

[b] Fold change was determined by using the normalized expression of *bolA* at each time point divided by the normalized expression of *bolA* in exponential phase.

Chapter 4. Discussion

To adapt to the stress and starved conditions in late stationary phase, *E. coli* cells mediate gene expression at the transcriptional level. Microarray experiments have been performed in the lab to characterize the gene expression from exponential phase to 48 h of incubation. However, due to the limitation of microarrays, such as cross hybridization between target genes and array probes (Chuaqui et al. 2002), microarrays may produce inaccurate results. In addition, the microarray data of gene expression at 24 and 48 h bacterial cultures was assayed from one biological replicate. The lack of biological replicate could have a great impact on the microarray results (Wurmbach et al. 2003). In this project, qPCR was performed to validate gene expression obtained from microarray experiments. Three genes with increased expression levels from 24 h to 48 h of incubation and two RpoS-dependent genes with decreased expression levels from 24h to 48 h of incubation were chosen to be the target genes. The identification of genes with increased expression levels from 24 h to 48 h of incubation will help elucidate how *E. coli* adapts to the long-term cultured environment. The functions of these genes, which include cell lysis, multidrug exporter, resistance to external stress and the alteration in morphology, are important for cell survival in late stationary phase. Two RpoS-dependent genes, *osmY* and *bolA*, with increased expression levels from exponential phase to early stationary phase were used as controls. The expression of these genes was quantified by the qPCR relative standard curve method.

4.1 RNA recovery in late stationary *E. coli*

Previously, traditional hot phenol-chloroform was successfully used to extract RNA from *E. coli* in exponential and early stationary phases. However, when RNA was extracted from *E. coli* in late stationary phase, the RNA recovery was found to be extremely low and the RNA was difficult to separate from genomic DNA. Although hot phenol-chloroform method was modified by adding DNase after first extraction to avoid extra extraction steps to remove DNase, RNA recovery at 24 h and 48 h of incubation was still very low (Table 3). Repeated extraction steps possibly contributed to the RNA loss during isolation. A commercial RNA extraction kit (Norgen Biotek, #17200) was also used to extract RNA. During extraction, 20 U of DNase was used to sufficiently remove genomic DNA (Table 4). The high genomic DNA contamination in 48 h bacterial cultures (Table 5) may be due to the accumulated DNA in large amounts of bacteria in late stationary phase. When cells are dead, DNA is released from lysed cells. The extracellular DNase produced by bacteria may not degrade all released DNA. Some evidence indicates that the released DNA can form macromolecular scaffolds in many different biofilms (Jakubovics et al. 2013), which supports the existence of free DNA in long-term cultured *E. coli*. The RNA loss during isolation may contribute to the low RNA recovery. To verify this hypothesis, the amount of PLYR (post lysis yield RNA) and recovered RNA was compared when RNA was extracted from *E. coli* in exponential phase and 48 h of incubation. However, the measured amount of PLYR (post lysis yield RNA) was even lower than that of recovered RNA (Table 6). The components of lysis solution possibly affected RNA concentration measurements. To evaluate the effect of

lysis solution on the measurement results, a standard curve was generated by measuring RNA concentration when RNA with the same concentration was resuspended in different percentages of lysis solutions. This standard curve implies that the lysis solution has a great effect on the measurement of RNA concentration. During isolation, RNA was resuspended in 75% lysis solution according to the instruction from Norgen Biotek. RNA was underestimated at least 40-fold when measured with the Qubit fluorometer (Figure 2). Thus, when RNA was extracted from cells of OD_{600} of 1.0 in exponential phase, the actual PLYR (post lysis yield RNA) amount would be 105 μg if we assumed 40-fold RNA was underestimated in 2648 ng RNA measured with the Qubit fluorometer (Table 6). So, 50% RNA was recovered when RNA was extracted from *E. coli* in exponential phase. 50% RNA was lost during isolation, which might be due to the limitation of the commercial kit (Norgen Biotek, #17200). Furthermore, when RNA was extracted from cells of OD_{600} of 1.0 at 48 h of incubation, the actual PLYR (post lysis yield RNA) was 5.28 μg. 28.4% RNA was recovered. The PLYR (post lysis yield RNA) amount was far less than 80 μg, which implies that the yield RNA amount could not be calculated by multiplying 10 pg (RNA amount per cell) with cell numbers. This is due to the large amounts of bacterial cells that begin to die when cells enter late stationary phase. RNA degradation also has a great impact on the estimation of yield RNA because RNA amount per cell changes in different growth phases (Mandelstam 1960).

4.2 Technical replication test by qPCR

Due to the limited size of the 96-well plate, there is not enough space to run three technical replicates at one time. The technical replication test of qPCR was performed prior to qPCR analysis. The expression of *tomB* at different time points was tested with one biological replicate in three technical replicates. The standard deviation was calculated to represent the variation of each technical operation. qPCR results partly confirmed the microarray data (Table 8). The discordance between these two methods might be due to the lack of biological replicates. In addition, qPCR and microarray has their inherent disadvantages. The accuracy of qPCR could be affected by amplification biases, primer dimers and the changing efficiency of qPCR at later cycles (Freeman et al. 1999). The non-specific cross hybridization between target genes and array probes could affect microarray results. Both methods are affected by the quality of RNA, because salts and alcohols in RNA samples could inhibit reverse transcription and PCR amplification (Freeman et al. 1999). The low expression levels of *rrsA* in 24 h cDNA samples may be caused by the low input amount of RNA for reverse transcription and qPCR. However, the function of *rrsA* was to normalize the difference among the RNA input quantity, thus the relative expression of *tomB* at different time points was not affected.

4.3 Quantification of target gene expression from exponential phase to 48 h of incubation

In this study, the expression of three genes, *hha*, *tomB* and *emrK* with increased expression levels from 24 h to 48 h of incubation obtained by microarray and two RpoS-

dependent genes, *osmY* and *bolA*, whose expression are highly induced in early stationary phase and repressed from 24 h to 48 h of incubation, was quantified by the qPCR relative standard curve method.

For the expression of *hha*, qPCR confirmed the increased expression from 24 h to 48 h of incubation. However, these two methods had different results for gene expression from exponential phase to early stationary phase. Microarray results indicated that the expression was repressed while qPCR results indicated that the expression levels stayed at the same level (Table 9). The low fold change (0.3) of microarray data between exponential phase and early stationary phase may result in this discordance. Previous study indicates that less than 2-fold change in microarray could produce a low correlation between microarray and qPCR results, especially in the validation of down-regulated genes (Morey et al. 2006) which agrees with the results. In late stationary phase, Hha can negatively regulate the expression of many genes with H-NS and then further result in the decreased metabolism (Hommais et al. 2001). In addition, Hha plays an important role in cell lysis and cell death (Wang et al. 2009a), which may explain the increased expression of *hha* over time. The increased expression of *tomB* from early stationary phase to 48 h of incubation was confirmed by qPCR (Table 10), which agreed with a previous finding that TomB and Hha can form a toxin and antitoxin module and is highly expressed in starvation (Garcia-Contreras et al. 2008). TomB can attenuate the toxic effect of Hha but cannot eliminate this toxin (Garcia-Contreras et al. 2008), which results in the overall decreased metabolism. The discordance about *tomB* expression from exponential phase to early stationary phase is possibly due to the low fold change (1) between exponential

phase and early stationary phase determined by microarray experiments (Table 10) (Morey et al. 2006). In addition, the low transcript abundance of *tomB* in exponential phase and early stationary phase (Figure 7, 8, Appendix A) increases the possibility of disagreement between microarray and qPCR results (Morey et al. 2006). qPCR results confirmed the increased expression of *emrK* from exponential phase to 48 h of incubation (Table 11). In late stationary phase, the synthesis of indole, a stationary phase extracellular signal, induces the expression of *emrK* (Hirakawa et al. 2005). The presence of H_2O_2 also induces the expression of *emrK* when cells grow under aerobic conditions (Han et al. 2010). The increased expression of *emrK* from exponential phase to 48 h of incubation supports these previous studies. The efflux function of *emrK* to export accumulated metabolites, such as quorum sensing molecules (Yang et al. 2005), is important for cell survival in long-term cultures. *bolA* and *osmY* are two RpoS-dependent genes. BolA plays an important role in cell morphology in stationary phase (Santos et al. 2002). The expression of *bolA* is induced in stationary phase and under stress conditions (Aldea et al. 1988). OsmY is involved in the response to hyperosmotic stress. The expression of *osmY* is also induced in stationary phase (Yim and Villarejo 1992). qPCR results partly confirmed microarray data that the expression of *osmY* and *bolA* increased in the early stationary phase. However, these two methods showed opposing results for the expression of *osmY* and *bolA* from early stationary to 24 h of incubation, and from 24 h to 48 h of incubation. qPCR results indicated that the expression of *osmY* decreased 17-fold and the expression of *bolA* decreased 3-fold from early stationary phase to 24 h of incubation. From 24 h to 48 h of incubation, the expression of *bolA* and *osmY* increased

around 3-fold (Table 12, 13). The different result for gene expression between 24 h to 48 h of incubation may be affected by the quality of RNA extracted from 48 h bacterial cultures. There was 9% genomic DNA contamination in these RNA samples (Table 5). The transcripts of *osmY* and *bolA*, especially *bolA*, were not abundant in the cells at 48 h of incubation (Figure 10, Appendix A), which increases the possibility of disagreeing results between microarray and qPCR experiments (Morey et al. 2006). In 48 h samples, the Cq difference of *bolA* and *osmY* between cDNA and NRT control was around 3 (1 Cq difference indicates around 2-fold difference in initial template input), which indicated 12.5% (1:8) amplification that could be ascribed to the genomic DNA contamination.

qPCR confirmed the directional change (up or down) between each growth phase. The degree of fold changes varied between these two methods, which agreed with a previous study that the fold changes determined by qPCR are significantly greater than fold changes assayed from microarray. The degree of fold change between these two methods, at least partly, depends on the platform and normalization methods of microarray (Yuen et al. 2002).

Chapter 5. Conclusion and future work

E. coli cells in late stationary phase undergo many physiological changes to adapt to the starved and stress conditions, such as mediating the expression of many genes at transcriptional level. In this study, qPCR was performed to quantify gene expression from exponential phase to 48 h of incubation and the results were compared with previous

microarray results. Genes with increased expression level between 24 h and 48 h bacterial cultures, including *hha, tomB* and *emrK,* as well as two RpoS-dependent genes with increased expression in early stationary phase and decreased expression from 24 h to 48 h of incubation, including *bolA* and *osmY,* were chosen as the target genes. qPCR results partly confirmed the microarray results. The expression of *hha* increases from 24 h to 48 h of incubation. The expression of *tomB* increases from early stationary phase to 48 h of incubation. The expression of *emrK* keeps increasing from exponential phase to 48 h of incubation. The discordant results between qPCR and microarray may be caused by the genomic DNA contamination of RNA samples isolated from 48 h cultures and the low abundance transcripts of target genes. Low fold changes between growth phases in the microarray analysis may also contribute to the differing results. The low transcript abundance of target genes increases the possibility of disagreeing results between these two methods. 20 U of DNase is sufficient to remove genomic DNA in exponential phase, early stationary phase and 24 h of incubation when RNA is extracted using the commercial kit (Norgen Biotek). RNA recovery decreases over time, which may be due to RNA loss during RNA isolation and cell death.

In the future, other genes with increased expression levels from 24 h to 48 h of incubation will be quantified at different time points by the relative standard curve method of qPCR. RNA-Seq results will be analyzed and used to validate the microarray and qPCR results. rRNA was not depleted before sequencing, so more genes encoding rRNA could be chosen as the reference gene for further qPCR experiments. *lacZ* fusion experiments can also be performed to validate gene expression in *E. coli* in late stationary

phase. The identification of genes expression in *E. coli* in late stationary phase will help better understand how *E. coli* adapts to long-term cultures and eventually help people control bacterial contamination in the environment.

References

Akerlund T, Nordstrom K, Bernander R (1995) Analysis of cell size and DNA content in exponentially growing and stationary-phase batch cultures of *Escherichia coli*. J Bacteriol 177:6791

Aldea M, Hernandez-Chico C, de la Campa AG, Kushner SR, Vicente M (1988) Identification, cloning, and expression of *bolA*, an *ftsZ*-dependent morphogene of *Escherichia coli*. J Bacteriol 170:5169

Allen RJ, Scott GK (1979) Biosynthesis and turnover of outer-membrane proteins in *Escherichia coli* ML308-225. Biochem J 182:407

Altuvia S, Almiron M, Huisman G, Kolter R, Storz G (1994) The *dps* promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. Mol Microbiol 13:265

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol 55:373

Banos RC, Vivero A, Aznar S, Garcia J, Pons M, Madrid C, Juarez A (2009) Differential regulation of horizontally acquired and core genome genes by the bacterial modulator H-NS. PLoS Genet 5:e1000513

Barker MM, Gaal T, Gourse RL (2001) Mechanism of regulation of transcription initiation by ppGpp. II. Models for positive control based on properties of RNAP mutants and competition for RNAP. J Mol Biol 305:689

Barrios AF, Zuo R, Ren D, Wood TK (2006) Hha, YbaJ, and OmpA regulate *Escherichia coli* K12 biofilm formation and conjugation plasmids abolish motility. Biotechnol Bioeng 93:188

Battesti A, Bouveret E (2006) Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. Mol Microbiol 62:1048

Bianchi AA, Baneyx F (1999) Hyperosmotic shock induces the sigma32 and sigmaE stress regulons of *Escherichia coli*. Mol Microbiol 34:1029

Braun V, Mahren S, Ogierman M (2003) Regulation of the FecI-type ECF sigma factor by transmembrane signalling. Curr Opin Microbiol 6:173

Brinkman AB, Ettema TJ, de Vos WM, van der Oost J (2003) The Lrp family of transcriptional regulators. Mol Microbiol 48:287

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611

Buts L, Lah J, Dao-Thi MH, Wyns L, Loris R (2005) Toxin-antitoxin modules as bacterial metabolic stress managers. Trends Biochem Sci 30:672

Castanie-Cornet MP, Penfound TA, Smith D, Elliott JF, Foster JW (1999) Control of acid resistance in *Escherichia coli*. J Bacteriol 181:3525

Chapman AG, Fall L, Atkinson DE (1971) Adenylate energy charge in *Escherichia coli* during growth and starvation. J Bacteriol 108:1072

Chiang SM, Schellhorn HE (2012) Regulators of oxidative stress response genes in *Escherichia coli* and their functional conservation in bacteria. Arch Biochem Biophys 525:161

Christensen SK, Mikkelsen M, Pedersen K, Gerdes K (2001) RelE, a global inhibitor of translation, is activated during nutritional stress. Proc Natl Acad Sci U S A 98:14328

Chuaqui RF, Bonner RF, Best CJ, Gillespie JW, Flaig MJ, Hewitt SM, Phillips JL, Krizman DB, Tangrea MA, Ahram M, Linehan WM, Knezevic V, Emmert-Buck MR (2002) Post-analysis follow-up and validation of microarray experiments. Nat Genet 32 Suppl:509

Colland F, Barth M, Hengge-Aronis R, Kolb A (2000) sigma factor selectivity of *Escherichia coli* RNA polymerase: role for CRP, IHF and lrp transcription factors. EMBO J 19:3028

Cronan JE, Jr. (1968) Phospholipid alterations during growth of *Escherichia coli.* J Bacteriol 95:2054

Davis BD, Luger SM, Tai PC (1986) Role of ribosome degradation in the death of starved *Escherichia coli* cells. J Bacteriol 166:439

Deutscher MP (2003) Degradation of stable RNA in bacteria. J Biol Chem 278:45041

Dhavan GM, Crothers DM, Chance MR, Brenowitz M (2002) Concerted binding and bending of DNA by *Escherichia coli* integration host factor. J Mol Biol 315:1027

Ding Q, Kusano S, Villarejo M, Ishihama A (1995) Promoter selectivity control of *Escherichia coli* RNA polymerase by ionic strength: differential recognition of osmoregulated promoters by E sigma D and E sigma S holoenzymes. Mol Microbiol 16:649

Ditto MD, Roberts D, Weisberg RA (1994) Growth phase variation of integration host factor level in *Escherichia coli*. J Bacteriol 176:3738

Dong T, Kirchhof MG, Schellhorn HE (2008) RpoS regulation of gene expression during exponential growth of *Escherichia coli* K12. Mol Genet Genomics 279:267

Dorman CJ, Barr GC, Ni Bhriain N, Higgins CF (1988) DNA supercoiling and the anaerobic and growth phase regulation of *tonB* gene expression. J Bacteriol 170:2816

Dukan S, Nystrom T (1998) Bacterial senescence: stasis results in increased and differential oxidation of cytoplasmic proteins leading to developmental induction of the heat shock regulon. Genes Dev 12:3431

Duquesne S, Destoumieux-Garzon D, Peduzzi J, Rebuffat S (2007) Microcins, gene-encoded antibacterial peptides from enterobacteria. Nat Prod Rep 24:708

Engelberg-Kulka H, Amitai S, Kolodkin-Gal I, Hazan R (2006) Bacterial programmed cell death and multicellular behavior in bacteria. PLoS Genet 2:e135

Ernsting BR, Atkinson MR, Ninfa AJ, Matthews RG (1992) Characterization of the regulon controlled by the leucine-responsive regulatory protein in *Escherichia coli*. J Bacteriol 174:1109

Falker S, Schmidt MA, Heusipp G (2006) Altered Ca(2+) regulation of Yop secretion in *Yersinia enterocolitica* after DNA adenine methyltransferase overproduction is mediated by Clp-dependent degradation of LcrG. J Bacteriol 188:7072

Farewell A, Diez AA, DiRusso CC, Nystrom T (1996) Role of the *Escherichia coli* FadR regulator in stasis survival and growth phase-dependent expression of the *uspA*, *fad*, and *fab* genes. J Bacteriol 178:6443

Farewell A, Kvint K, Nystrom T (1998) Negative regulation by RpoS: a case of sigma factor competition. Mol Microbiol 29:1039

Ferenci T (2005) Maintaining a healthy SPANC balance through regulatory and mutational adaptation. Mol Microbiol 57:1

Finkel SE (2006) Long-term survival during stationary phase: evolution and the GASP phenotype. Nat Rev Microbiol 4:113

Freeman WM, Walker SJ, Vrana KE (1999) Quantitative RT-PCR: pitfalls and potential. Biotechniques 26:112

Fu J, Li D, Xia S, Song H, Dong Z, Chen F, Sun X, Tang Z (2009) Absolute quantification of plasmid DNA by real-time PCR with genomic DNA as external standard and its application to a biodistribution study of an HIV DNA vaccine. Anal Sci 25:675

Garcia-Contreras R, Zhang XS, Kim Y, Wood TK (2008) Protein translation and cell death: the role of rare tRNAs in biofilm formation and in activating dormant phage killer genes. PLoS One 3:e2394

Gentry DR, Hernandez VJ, Nguyen LH, Jensen DB, Cashel M (1993) Synthesis of the stationary-phase sigma factor sigma s is positively regulated by ppGpp. J Bacteriol 175:7982

Giannoukos G, Ciulla DM, Huang K, Haas BJ, Izard J, Levin JZ, Livny J, Earl AM, Gevers D, Ward DV, Nusbaum C, Birren BW, Gnirke A (2012) Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes. Genome Biol 13:R23

Gonzalez-Flecha B, Demple B (1997) Homeostatic regulation of intracellular hydrogen peroxide concentration in aerobically growing *Escherichia coli*. J Bacteriol 179:382

Groat RG, Schultz JE, Zychlinsky E, Bockman A, Matin A (1986) Starvation proteins in *Escherichia coli:* kinetics of synthesis and role in starvation survival. J Bacteriol 168:486

Grossman AD, Erickson JW, Gross CA (1984) The htpR gene product of *E. coli* is a sigma factor for heat-shock promoters. Cell 38:383

Gruber TM, Gross CA (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. Annu Rev Microbiol 57:441

Haas BJ, Chin M, Nusbaum C, Birren BW, Livny J (2012) How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? BMC Genomics 13:734

Hall BG (1991) Adaptive evolution that requires multiple spontaneous mutations: mutations involving base substitutions. Proc Natl Acad Sci U S A 88:5882

Han X, Dorsey-Oresto A, Malik M, Wang JY, Drlica K, Zhao X, Lu T (2010) *Escherichia coli* genes that reduce the lethal effects of stress. BMC Microbiol 10:35

Hartley CL, Clements HM, Linton KB (1977) *Escherichia coli* in the faecal flora of man. J Appl Bacteriol 43:261

Haseltine WA, Block R (1973) Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes. Proc Natl Acad Sci U S A 70:1564

Helmann JD (1991) Alternative sigma factors and the regulation of flagellar gene expression. Mol Microbiol 5:2875

Hengge-Aronis R, Klein W, Lange R, Rimmele M, Boos W (1991) Trehalose synthesis genes are controlled by the putative sigma factor encoded by *rpoS* and are involved in stationary-phase thermotolerance in *Escherichia coli*. J Bacteriol 173:7918

Hirakawa H, Inazumi Y, Masaki T, Hirata T, Yamaguchi A (2005) Indole induces the expression of multidrug exporter genes in *Escherichia coli*. Mol Microbiol 55:1113

Hirschman J, Wong PK, Sei K, Keener J, Kustu S (1985) Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription in vitro: evidence that the *ntrA* product is a sigma factor. Proc Natl Acad Sci U S A 82:7525

Hommais F, Krin E, Laurent-Winter C, Soutourina O, Malpertuy A, Le Caer JP, Danchin A, Bertin P (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. Mol Microbiol 40:20

Hunt SM, Werner EM, Huang B, Hamilton MA, Stewart PS (2004) Hypothesis for the role of nutrient starvation in biofilm detachment. Appl Environ Microbiol 70:7418

Imlay KR, Imlay JA (1996) Cloning and analysis of *sodC,* encoding the copper-zinc superoxide dismutase of Escherichia coli. J Bacteriol 178:2564

Ishihama A (2000) Functional modulation of *Escherichia coli* RNA polymerase. Annu Rev Microbiol 54:499

Izutsu K, Wada A, Wada C (2001) Expression of ribosome modulation factor (RMF) in *Escherichia coli* requires ppGpp. Genes Cells 6:665

Jakubovics NS, Shields RC, Rajarajan N, Burgess JG (2013) Life after death: the critical role of extracellular DNA in microbial biofilms. Lett Appl Microbiol

Jenkins DE, Chaisson SA, Matin A (1990) Starvation-induced cross protection against osmotic challenge in *Escherichia coli*. J Bacteriol 172:2779

Jenkins DE, Schultz JE, Matin A (1988) Starvation-induced cross protection against heat or H_2O_2 challenge in *Escherichia coli*. J Bacteriol 170:3910

Jishage M, Ishihama A (1998) A stationary phase protein in *Escherichia coli* with binding activity to the major sigma subunit of RNA polymerase. Proc Natl Acad Sci U S A 95:4953

Jishage M, Iwata A, Ueda S, Ishihama A (1996) Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of four species of sigma subunit under various growth conditions. J Bacteriol 178:5447

Jishage M, Kvint K, Shingler V, Nystrom T (2002) Regulation of sigma factor competition by the alarmone ppGpp. Genes Dev 16:1260

Justino MC, Almeida CC, Teixeira M, Saraiva LM (2007) *Escherichia coli* di-iron YtfE protein is necessary for the repair of stress-damaged iron-sulfur clusters. J Biol Chem 282:10352

Justino MC, Vicente JB, Teixeira M, Saraiva LM (2005) New genes implicated in the protection of anaerobically grown *Escherichia coli* against nitric oxide. J Biol Chem 280:2636

Kawano H, Hirokawa Y, Mori H (2009) Long-term survival of *Escherichia coli* lacking the HipBA toxin-antitoxin system during prolonged cultivation. Biosci Biotechnol Biochem 73:117

Keller L, Surette MG (2006) Communication in bacteria: an ecological and evolutionary perspective. Nat Rev Microbiol 4:249

Kirby JE, Trempy JE, Gottesman S (1994) Excision of a P4-like cryptic prophage leads to Alp protease expression in *Escherichia coli*. J Bacteriol 176:2068

Kobayashi A, Hirakawa H, Hirata T, Nishino K, Yamaguchi A (2006) Growth phase-dependent expression of drug exporters in *Escherichia coli* and its contribution to drug tolerance. J Bacteriol 188:5693

Kolter R, Siegele DA, Tormo A (1993) The stationary phase of the bacterial life cycle. Annu Rev Microbiol 47:855

Kuczynska-Wisnik D, Matuszewska E, Laskowska E (2010) *Escherichia coli* heat-shock proteins IbpA and IbpB affect biofilm formation by influencing the level of extracellular indole. Microbiology 156:148

Lange R, Barth M, Hengge-Aronis R (1993) Complex transcriptional control of the sigma s-dependent stationary-phase-induced and osmotically regulated *osmY* (csi-5) gene suggests novel roles for Lrp, cyclic AMP (cAMP) receptor protein-cAMP complex, and integration host factor in the stationary-phase response of *Escherichia coli*. J Bacteriol 175:7910

Lange R, Hengge-Aronis R (1991) Growth phase-regulated expression of *bolA* and morphology of stationary-phase *Escherichia coli* cells are controlled by the novel sigma factor sigma S. J Bacteriol 173:4474

Lazazzera BA (2000) Quorum sensing and starvation: signals for entry into stationary phase. Curr Opin Microbiol 3:177

Leland FE, Kohn DF, Sirbasku DA (1983) Effect of estrogen-promoted bacterial infections of the rat uterus on bioassay of mammalian cell growth factor activities in uterine luminal fluid. Biol Reprod 28:1243

Lewis K (2000) Programmed death in bacteria. Microbiol Mol Biol Rev 64:503

Lin J, Smith MP, Chapin KC, Baik HS, Bennett GN, Foster JW (1996) Mechanisms of acid resistance in enterohemorrhagic *Escherichia coli.* Appl Environ Microbiol 62:3094

Loewe L, Textor V, Scherer S (2003) High deleterious genomic mutation rate in stationary phase of *Escherichia coli.* Science 302:1558

Lonetto M, Gribskov M, Gross CA (1992) The sigma 70 family: sequence conservation and evolutionary relationships. J Bacteriol 174:3843

Lubelski J, Konings WN, Driessen AJ (2007) Distribution and physiology of ABC-type transporters contributing to multidrug resistance in bacteria. Microbiol Mol Biol Rev 71:463

Madrid C, Balsalobre C, Garcia J, Juarez A (2007) The novel Hha/YmoA family of nucleoid-associated proteins: use of structural mimicry to modulate the activity of the H-NS family of proteins. Mol Microbiol 63:7

Maeda H, Fujita N, Ishihama A (2000) Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase. Nucleic Acids Res 28:3497

Magnusson LU, Farewell A, Nystrom T (2005) ppGpp: a global regulator in *Escherichia coli*. Trends Microbiol 13:236

Mandelstam J (1960) The intracellular turnover of protein and nucleic acids and its role in biochemical differentiation. Bacteriol Rev 24:289

Mangan MW, Lucchini S, Danino V, Croinin TO, Hinton JC, Dorman CJ (2006) The integration host factor (IHF) integrates stationary-phase and virulence gene expression in *Salmonella enterica serovar Typhimurium*. Mol Microbiol 59:1831

Marianovsky I, Aizenman E, Engelberg-Kulka H, Glaser G (2001) The regulation of the *Escherichia coli mazEF* promoter involves an unusual alternating palindrome. J Biol Chem 276:5975

Matsumura K, Furukawa S, Ogihara H, Morinaga Y (2011) Roles of multidrug efflux pumps on the biofilm formation of *Escherichia coli* K-12. Biocontrol Sci 16:69

Mecsas J, Rouviere PE, Erickson JW, Donohue TJ, Gross CA (1993) The activity of sigma E, an *Escherichia coli* heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. Genes Dev 7:2618

Mengin-Lecreulx D, van Heijenoort J (1985) Effect of growth conditions on peptidoglycan content and cytoplasmic steps of its biosynthesis in *Escherichia coli*. J Bacteriol 163:208

Merrick MJ (1993) In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N). Mol Microbiol 10:903

Miyadai H, Tanaka-Masuda K, Matsuyama S, Tokuda H (2004) Effects of lipoprotein overproduction on the induction of DegP (HtrA) involved in quality control in the *Escherichia coli* periplasm. J Biol Chem 279:39807

Morey JS, Ryan JC, Van Dolah FM (2006) Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. Biol Proced Online 8:175

Navarro Llorens JM, Tormo A, Martinez-Garcia E (2010) Stationary phase in gram-negative bacteria. FEMS Microbiol Rev 34:476

Nieto JM, Madrid C, Prenafeta A, Miquelay E, Balsalobre C, Carrascal M, Juarez A (2000) Expression of the hemolysin operon in *Escherichia coli* is modulated by a nucleoid-protein complex that includes the proteins Hha and H-NS. Mol Gen Genet 263:349

Nystrom T (1994) The glucose-starvation stimulon of *Escherichia coli*: induced and repressed synthesis of enzymes of central metabolic pathways and role of acetyl phosphate in gene expression and starvation survival. Mol Microbiol 12:833

Nystrom T (2004) Stationary-phase physiology. Annu Rev Microbiol 58:161

Nystrom T, Larsson C, Gustafsson L (1996) Bacterial defense against aging: role of the *Escherichia coli* ArcA regulator in gene expression, readjusted energy flux and survival during stasis. EMBO J 15:3219

Ohniwa RL, Morikawa K, Kim J, Ohta T, Ishihama A, Wada C, Takeyasu K (2006) Dynamic state of DNA topology is essential for genome condensation in bacteria. EMBO J 25:5591

Onufryk C, Crouch ML, Fang FC, Gross CA (2005) Characterization of six lipoproteins in the sigmaE regulon. J Bacteriol 187:4552

Paget MS, Helmann JD (2003) The sigma70 family of sigma factors. Genome Biol 4:203

Park S, Imlay JA (2003) High levels of intracellular cysteine promote oxidative DNA damage by driving the fenton reaction. J Bacteriol 185:1942

Patten CL, Kirchhof MG, Schertzberg MR, Morton RA, Schellhorn HE (2004) Microarray analysis of RpoS-mediated gene expression in *Escherichia coli* K-12. Mol Genet Genomics 272:580

Pieper U, Pingoud A (2002) A mutational analysis of the PD...D/EXK motif suggests that McrC harbors the catalytic center for DNA cleavage by the GTP-dependent restriction enzyme McrBC from Escherichia coli. Biochemistry 41:5236

Reeve CA, Amy PS, Matin A (1984) Role of protein synthesis in the survival of carbon-starved *Escherichia coli* K-12. J Bacteriol 160:1041

Ren D, Bedzyk LA, Thomas SM, Ye RW, Wood TK (2004) Gene expression in *Escherichia coli* biofilms. Appl Microbiol Biotechnol 64:515

Reyes-Dominguez Y, Contreras-Ferrat G, Ramirez-Santos J, Membrillo-Hernandez J, Gomez-Eichelmann MC (2003) Plasmid DNA supercoiling and gyrase activity in *Escherichia coli* wild-type and rpoS stationary-phase cells. J Bacteriol 185:1097

Robinow C, Kellenberger E (1994) The bacterial nucleoid revisited. Microbiol Rev 58:211

Rolfe MD, Rice CJ, Lucchini S, Pin C, Thompson A, Cameron AD, Alston M, Stringer MF, Betts RP, Baranyi J, Peck MW, Hinton JC (2012) Lag phase is a distinct

growth phase that prepares bacteria for exponential growth and involves transient metal accumulation. J Bacteriol 194:686

Ross TK, Achberger EC, Braymer HD (1989) Nucleotide sequence of the McrB region of *Escherichia coli* K-12 and evidence for two independent translational initiation sites at the mcrB locus. J Bacteriol 171:1974

Rutishauser J, Boni-Schnetzler M, Boni J, Wichmann W, Huisman T, Vallotton MB, Froesch ER (1996) A novel point mutation in the translation initiation codon of the pre-pro-vasopressin-neurophysin II gene: cosegregation with morphological abnormalities and clinical symptoms in autosomal dominant neurohypophyseal diabetes insipidus. J Clin Endocrinol Metab 81:192

Saint-Ruf C, Matic I (2006) Environmental tuning of mutation rates. Environ Microbiol 8:193

Santos JM, Lobo M, Matos AP, De Pedro MA, Arraiano CM (2002) The gene *bolA* regulates *dacA* (PBP5), *dacC* (PBP6) and *ampC* (AmpC), promoting normal morphology in *Escherichia coli.* Mol Microbiol 45:1729

Sasagawa Y, Nikaido I, Hayashi T, Danno H, Uno KD, Imai T, Ueda HR (2013) Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol 14:R31

Schellhorn HE (1995) Regulation of hydroperoxidase (catalase) expression in *Escherichia coli.* FEMS Microbiol Lett 131:113

Schellhorn HE, Hassan HM (1988) Transcriptional regulation of *katE* in *Escherichia coli* K-12. J Bacteriol 170:4286

Seyfzadeh M, Keener J, Nomura M (1993) *spoT*-dependent accumulation of guanosine tetraphosphate in response to fatty acid starvation in *Escherichia coli*. Proc Natl Acad Sci U S A 90:11004

Sezonov G, Joseleau-Petit D, D'Ari R (2007) *Escherichia coli* physiology in Luria-Bertani broth. J Bacteriol 189:8746

Shah D, Zhang Z, Khodursky A, Kaldalu N, Kurg K, Lewis K (2006) Persisters: a distinct physiological state of *E. coli.* BMC Microbiol 6:53

Sharma VK, Zuerner RL (2004) Role of hha and ler in transcriptional regulation of the esp operon of enterohemorrhagic *Escherichia coli* O157:H7. J Bacteriol 186:7290

Snell EE (1975) Tryptophanase: structure, catalytic activities, and mechanism of action. Adv Enzymol Relat Areas Mol Biol 42:287

Soto SM (2013) Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm. Virulence 4:223

Sperandio B, Polard P, Ehrlich DS, Renault P, Guedon E (2005) Sulfur amino acid metabolism and its control in *Lactococcus lactis* IL1403. J Bacteriol 187:3762

Spira B, Silberstein N, Yagil E (1995) Guanosine 3',5'-bispyrophosphate (ppGpp) synthesis in cells of *Escherichia coli* starved for Pi. J Bacteriol 177:4053

Strauch KL, Johnson K, Beckwith J (1989) Characterization of *degP*, a gene required for proteolysis in the cell envelope and essential for growth of *Escherichia coli* at high temperature. J Bacteriol 171:2689

Surette MG, Bassler BL (1998) Quorum sensing in *Escherichia coli* and *Salmonella typhimurium.* Proc Natl Acad Sci U S A 95:7046

Tani TH, Khodursky A, Blumenthal RM, Brown PO, Matthews RG (2002) Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. Proc Natl Acad Sci U S A 99:13471

Todorovic S, Justino MC, Wellenreuther G, Hildebrandt P, Murgida DH, Meyer-Klaucke W, Saraiva LM (2008) Iron-sulfur repair YtfE protein from *Escherichia coli*: structural characterization of the di-iron center. J Biol Inorg Chem 13:765

Vieira HL, Freire P, Arraiano CM (2004) Effect of *Escherichia coli* morphogene *bolA* on biofilms. Appl Environ Microbiol 70:5682

Vijayakumar SR, Kirchhof MG, Patten CL, Schellhorn HE (2004) RpoS-regulated genes of *Escherichia coli* identified by random *lacZ* fusion mutagenesis. J Bacteriol 186:8499

Vinella D, Albrecht C, Cashel M, D'Ari R (2005) Iron limitation induces SpoT-dependent accumulation of ppGpp in *Escherichia coli.* Mol Microbiol 56:958

Wada A, Igarashi K, Yoshimura S, Aimoto S, Ishihama A (1995) Ribosome modulation factor: stationary growth phase-specific inhibitor of ribosome functions from *Escherichia coli.* Biochem Biophys Res Commun 214:410

Wang X, Kim Y, Ma Q, Hong SH, Pokusaeva K, Sturino JM, Wood TK (2010) Cryptic prophages help bacteria cope with adverse environments. Nat Commun 1:147

Wang X, Kim Y, Wood TK (2009a) Control and benefits of CP4-57 prophage excision in *Escherichia coli* biofilms. ISME J 3:1164

Wang Z, Gerstein M, Snyder M (2009b) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57

Weber H, Polen T, Heuveling J, Wendisch VF, Hengge R (2005) Genome-wide analysis of the general stress response network in *Escherichia coli*: sigmaS-dependent genes, promoters, and sigma factor selectivity. J Bacteriol 187:1591

Weichart D, Querfurth N, Dreger M, Hengge-Aronis R (2003) Global role for ClpP-containing proteases in stationary-phase adaptation of Escherichia coli. J Bacteriol 185:115

Wion D, Casadesus J (2006) N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. Nat Rev Microbiol 4:183

Wolf SG, Frenkiel D, Arad T, Finkel SE, Kolter R, Minsky A (1999) DNA protection by stress-induced biocrystallization. Nature 400:83

Wurmbach E, Yuen T, Sealfon SC (2003) Focused microarray analysis. Methods 31:306

Xiao H, Kalman M, Ikehara K, Zemel S, Glaser G, Cashel M (1991) Residual guanosine 3',5'-bispyrophosphate synthetic activity of *relA* null mutants can be eliminated by *spoT* null mutations. J Biol Chem 266:5980

Yang F, Wang LH, Wang J, Dong YH, Hu JY, Zhang LH (2005) Quorum quenching enzyme activity is widely conserved in the sera of mammalian species. FEBS Lett 579:3713

Yim HH, Villarejo M (1992) *osmY*, a new hyperosmotically inducible gene, encodes a periplasmic protein in *Escherichia coli*. J Bacteriol 174:3637

Yuen T, Wurmbach E, Pfeffer RL, Ebersole BJ, Sealfon SC (2002) Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. Nucleic Acids Res 30:e48

Zhao K, Liu M, Burgess RR (2010) Promoter and regulon analysis of nitrogen assimilation factor, sigma54, reveal alternative strategy for *E. coli* MG1655 flagellar biosynthesis. Nucleic Acids Res 38:1273

Zhou K, Zhou L, Lim Q, Zou R, Stephanopoulos G, Too HP (2011) Novel reference genes for quantifying transcriptional responses of *Escherichia coli* to protein overexpression by quantitative PCR. BMC Mol Biol 12:18

Appendix A.

Table 1.The expression of target genes from exponential phase to 48 h of incubation assayed by microarray.

The expression intensities of target genes were $\log_2$ transformed. The microarray data of gene expression in exponential phase and early stationary phase was obtained from Dong (Dong et al. 2008). The microarray data of gene expression in 24 h and 48 h of incubation was obtained from S. Chiang (unpublished). Fold change was used to represent the expression change between each growth phase with exponential phase.

Time	Log2 of expression intensity					Fold change				
(h)	*hha*	*tomB*	*emrK*	*bolA*	*osmY*	*hha*	*tomB*	*emrK*	*bolA*	*osmY*
5	8.6	5.0	2.8	3.2	6.7	1.0	1.0	1.0	1.0	1.0
7	6.6	5.1	3.3	7.6	10.5	0.3	1.0	1.4	20.0	14.0
24	8.2	7.2	4.3	11.7	12.4	1.0	5.0	2.7	345.0	52.0
48	11.1	9.6	6.6	4.7	7.9	7.8	24.0	13.1	3.0	8.0

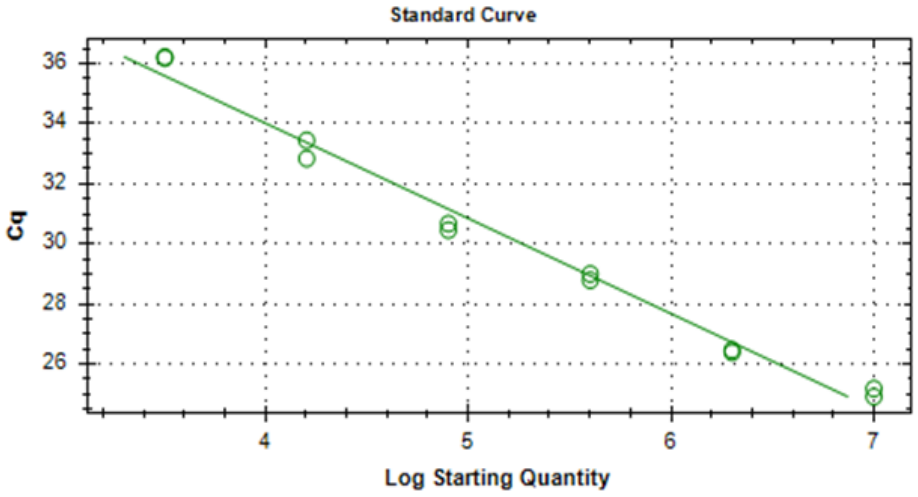


Figure 1. Amplification efficiency assay of *hha*. The standard curve was linear from 3.2 x 10^3 to 2.0 x 10^6 gene copies. The amplification efficiency (E) was 106.7%. The coefficient of determination (R^2) was 0.985. The quantitative cycle (Cq) for the standard curve ranged from 25 to 35.This graph was exported from Bio-Rad CFX Manager™.

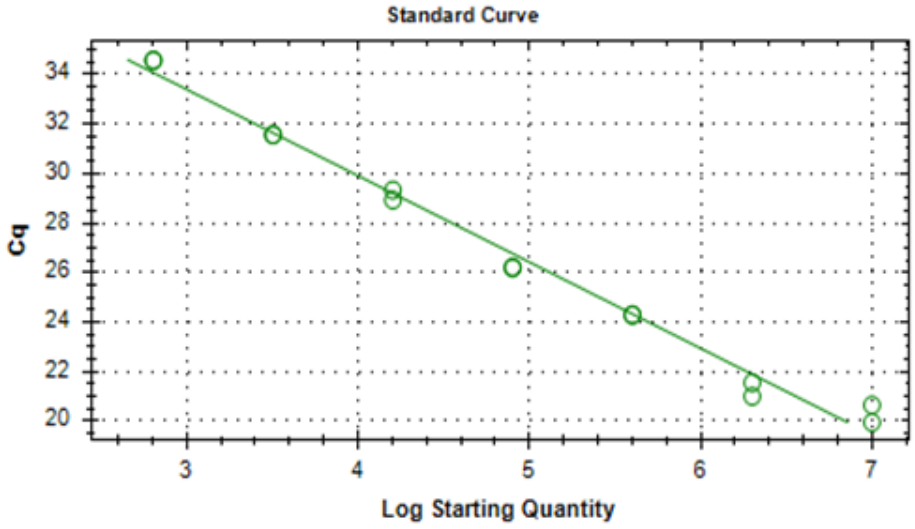


Figure 2. **Amplification efficiency assay of *rrsA*.** The standard curve was linear from 6.4 x 10^2 to 2.0 x 10^6 gene copies. The amplification efficiency (E) was 93.6% and the coefficient of determination (R^2) was 0.989. The quantitative cycle (Cq) for the standard curve ranged from 20 to 35. This graph was exported from Bio-Rad CFX Manager™.

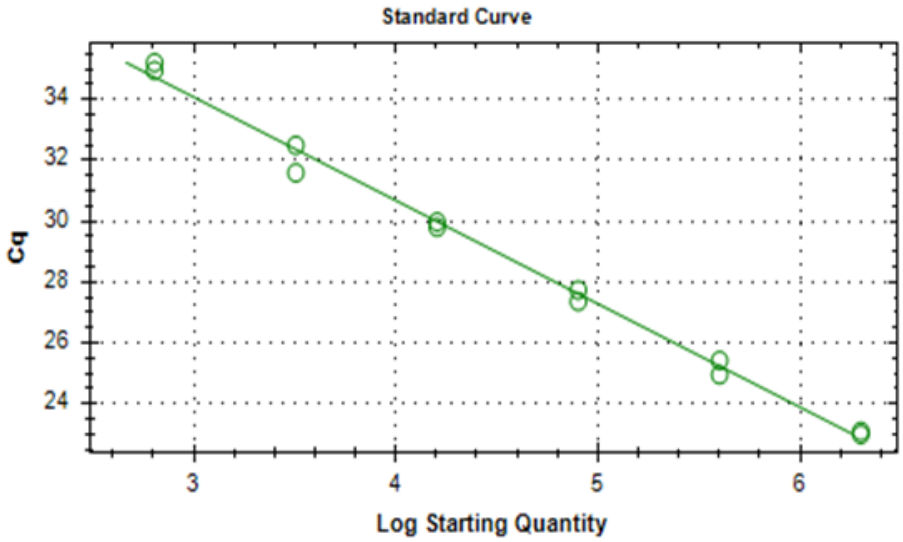


Figure 3. Amplification efficiency assay of *tomB*. The standard curve was linear from 3.2 x 10^3 to 2.0 x 10^6 gene copies. The amplification efficiency (E) was 97% and the coefficient of determination (R^2) was 0.994. The quantitative cycle (Cq) for the standard curve ranged from 23 to 35. This graph was exported from Bio-Rad CFX Manager™.

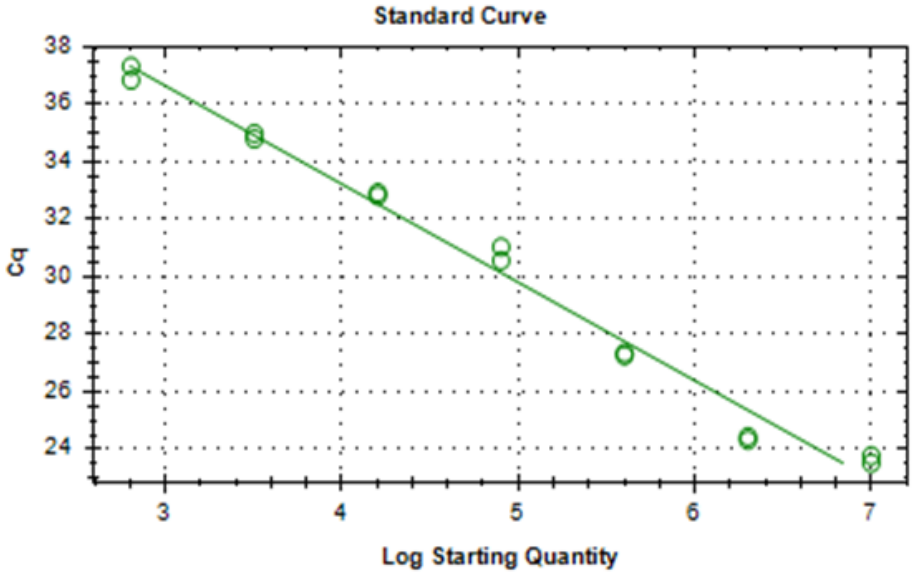


Figure 4. Amplification efficiency assay of *emrK*. The standard curve was linear from 6.4 x 10^2 to 2.0 x 10^6 gene copies. The amplification efficiency (E) was 96.1% and the coefficient of determination (R^2) was 0.985.The quantitative cycle (Cq) for the standard curve ranged from 23 to 37. This graph was exported from Bio-Rad CFX Manager™.

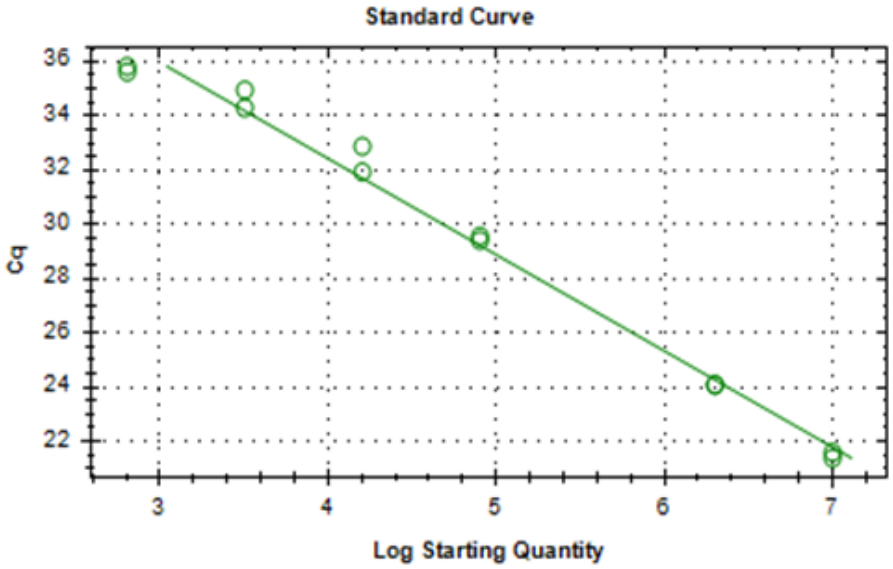


Figure 5. Amplification efficiency assay of *bolA*. The standard curve was linear from 6.4 x 10^2 to 2.0 x 10^6 gene copies. The amplification efficiency (E) was 91.3% and the coefficient of determination (R^2) was 0.987.The quantitative cycle (Cq) for the standard curve ranged from 21 to 35. This graph was exported from Bio-Rad CFX Manager™.

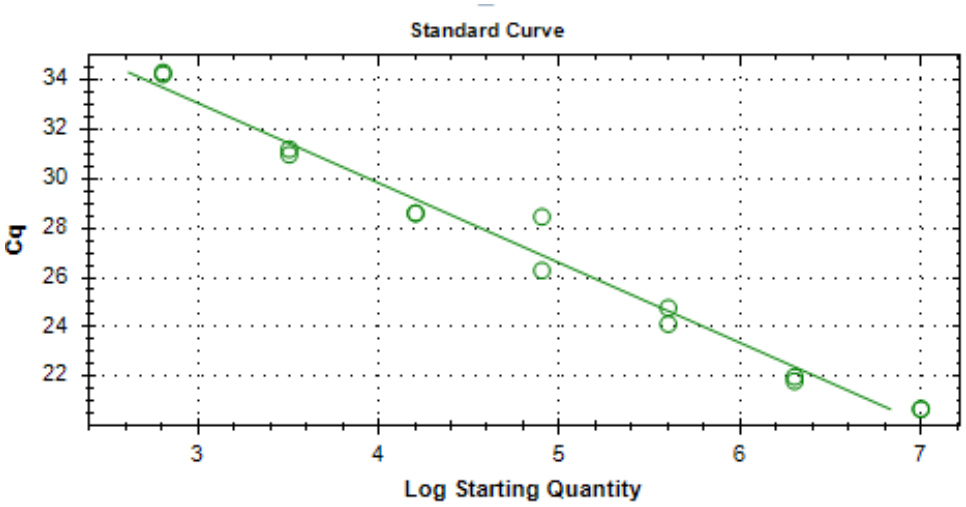


Figure 6. Amplification efficiency assay of *osmY*. The standard curve was linear from 6.4 x 10^2 to 2.0 x 10^6 gene copies. The amplification efficiency (E) was 101.5% and the coefficient of determination (R^2) was 0.980.The quantitative cycle (Cq) for the standard curve ranged from 20 to 34. This graph was exported from Bio-Rad CFX Manager™.

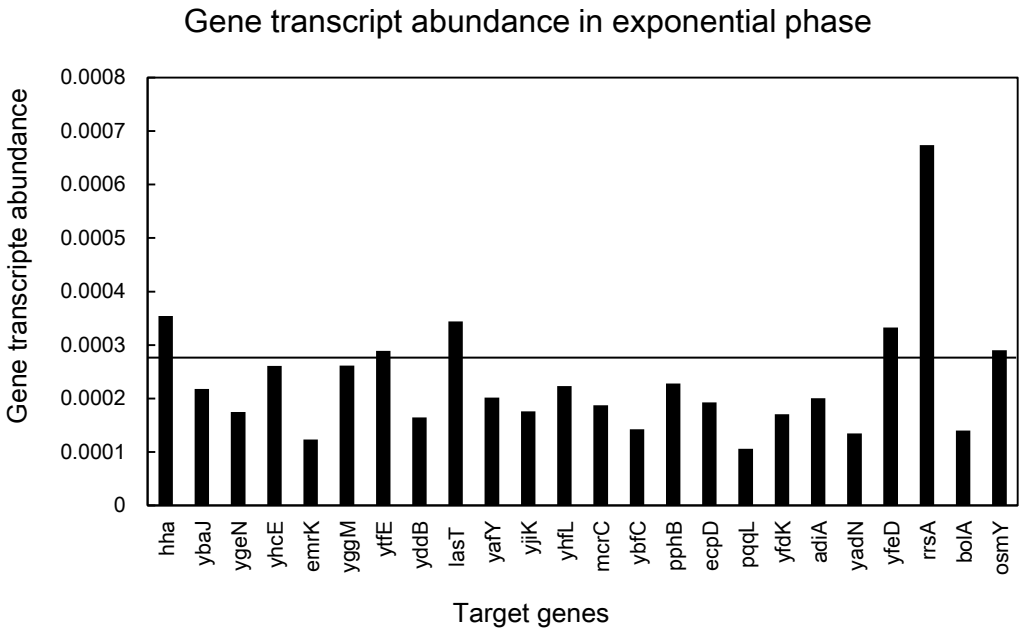


Figure 7. Gene transcript abundance in exponential phase. The average of gene transcripts was 0.000228, which was calculated by dividing the expression intensities of total genes with total gene numbers. 20 genes with highly increased expression levels and two genes (*bolA* and *osmY*) with decreased expression levels from 24 h to 48 h of incubation determined by microarray as well as *rrsA* were chosen to be the target genes. The transcript abundance of target gene was calculated by dividing the expression intensity of each gene with the expression intensities of total genes obtained by microarray.

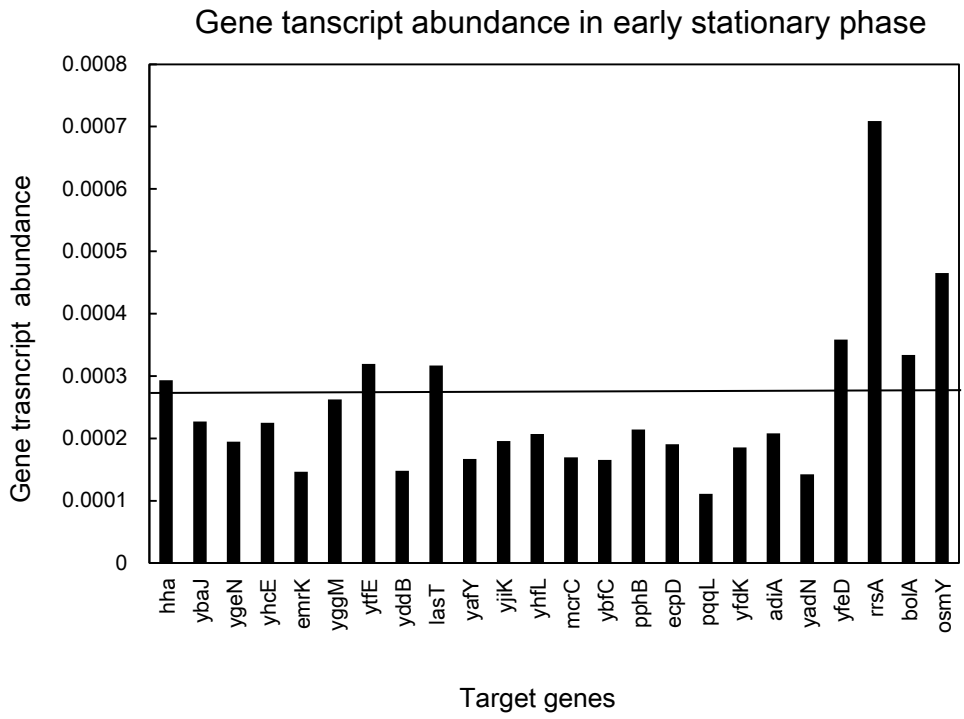


Figure 8. Gene transcript abundance in early stationary phase. The average of gene transcripts was 0.000228, which was calculated by dividing the expression intensities of total genes with total gene numbers. 20 genes with highly increased expression levels and two genes (*bolA* and *osmY*) with decreased expression levels from 24 h to 48 h of incubation determined by microarray as well as *rrsA* were chosen to be the target genes. The transcript abundance of target gene was calculated by dividing the expression intensity of each gene with the expression intensities of total genes obtained by microarray.

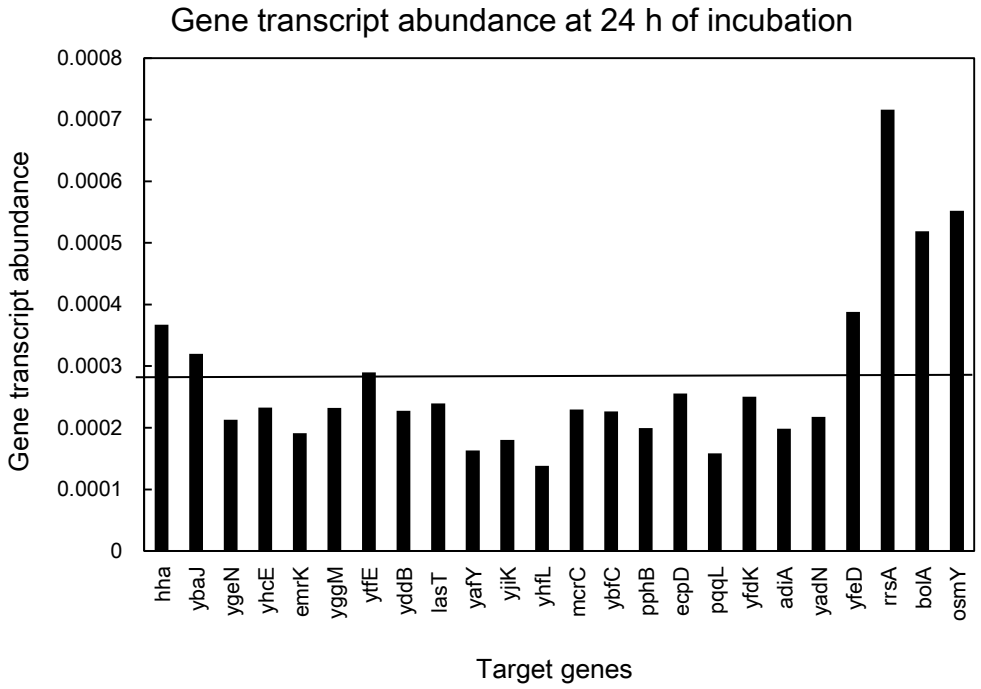


Figure 9. Gene transcript abundance at 24 h of incubation. The average of gene transcripts was 0.000228, which was calculated by dividing the expression intensities of total genes with total gene numbers. 20 genes with highly increased expression levels and two genes (*bolA* and *osmY*) with decreased expression levels from 24 h to 48 h of incubation determined by microarray as well as *rrsA* were chosen to be the target genes. The transcript abundance of target gene was calculated by dividing the expression intensity of each gene with the expression intensities of total genes obtained by microarray.

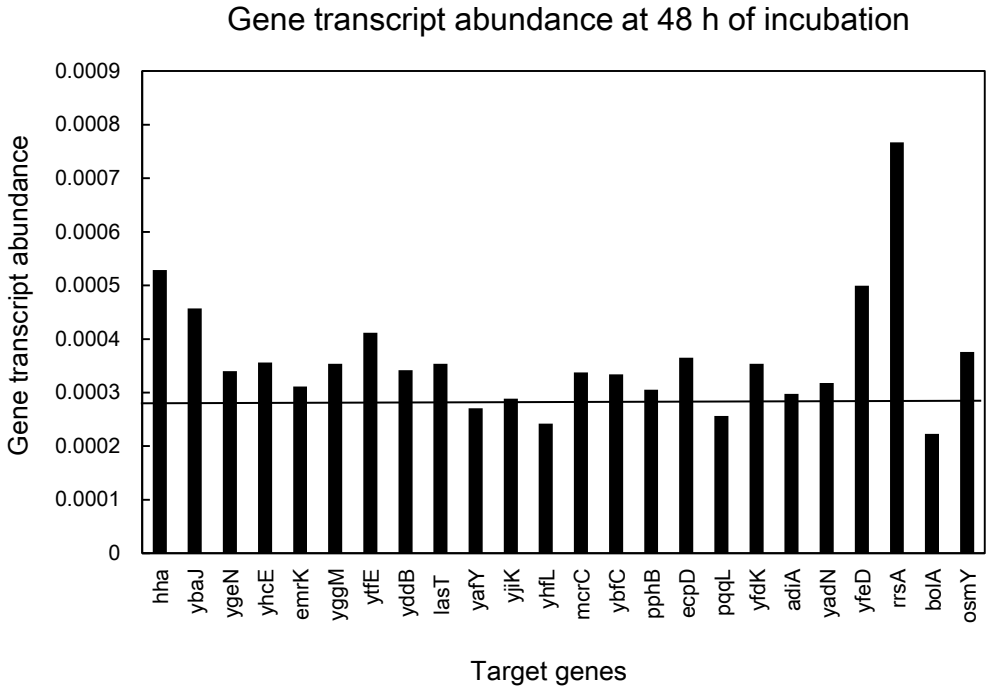


Figure 10. Gene transcript abundance at 48 h of incubation. The average of gene transcripts was 0.000228, which was calculated by dividing the expression intensities of total genes with total gene numbers. 20 genes with highly increased expression levels and two genes (*bolA* and *osmY*) with decreased expression levels from 24 h to 48 h of incubation determined by microarray as well as *rrsA* were chosen to be the target genes. The transcript abundance of target gene was calculated by dividing the expression intensity of each gene with the expression intensities of total genes obtained by microarray.

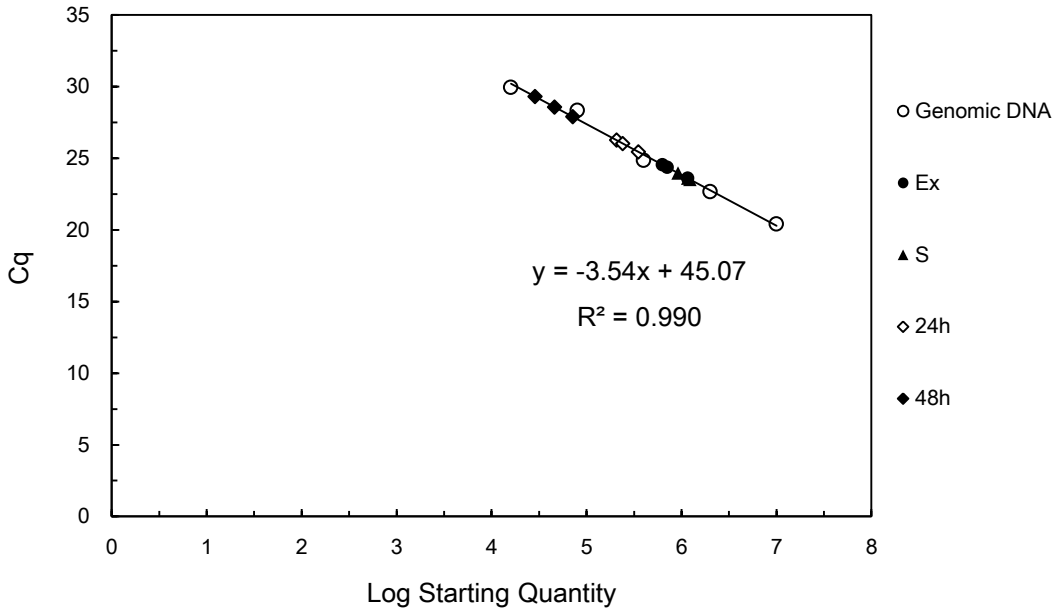


Figure 11. *rrsA* Standard curve and Cq values of cDNA from exponential phase to 48 h of incubation. White circle symbols indicate the Cq values of the serially diluted genomic DNA which were used as the templates to construct standard curve. The black cycles indicate the Cq values of cDNA in exponential phase. The black triangles indicate Cq values of cDNA in early stationary phase. The white diamonds indicate Cq values of cDNA at 24 h of incubation. The black diamonds indicate the Cq values of cDNA at 48 h of incubation. The amplification efficiency (E) was 91.6% and the coefficient of determination (R^2) was 0.990. The dynamic range was between 1.6 x 10^4 and 4.0 x 10^5 gene copies. The Cq values of tested cDNA samples were within the dynamic range.

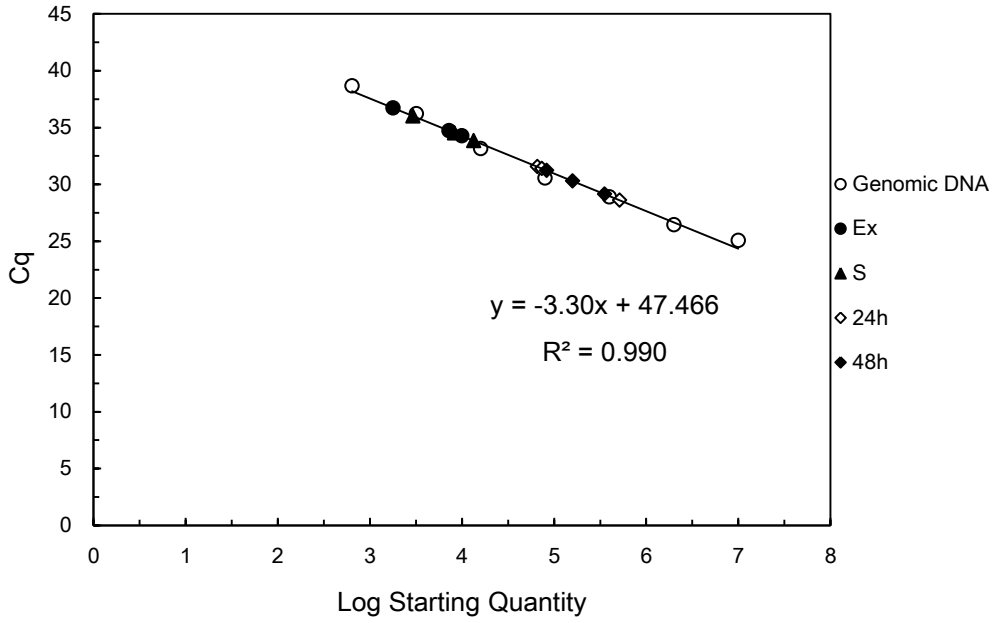


Figure 12. *hha* Standard curve and Cq values of cDNA from exponential phase to 48h of incubation. White circle symbols indicate the Cq values of the serially diluted genomic DNA which were used as the templates to construct standard curve. The black cycles indicate the Cq values of cDNA in exponential phase. The black triangles indicate Cq values of cDNA in early stationary phase. The white diamonds indicate Cq values of cDNA at 24 h of incubation. The black diamonds indicate the Cq values of cDNA at 48 h of incubation. The amplification efficiency (E) was 100.9% and the coefficient of determination (R^2) was 0.990. The dynamic range was between 6.4 x 10^2 and 1.0 x 10^7 gene copies. The Cq values of tested cDNA samples were within the dynamic range.

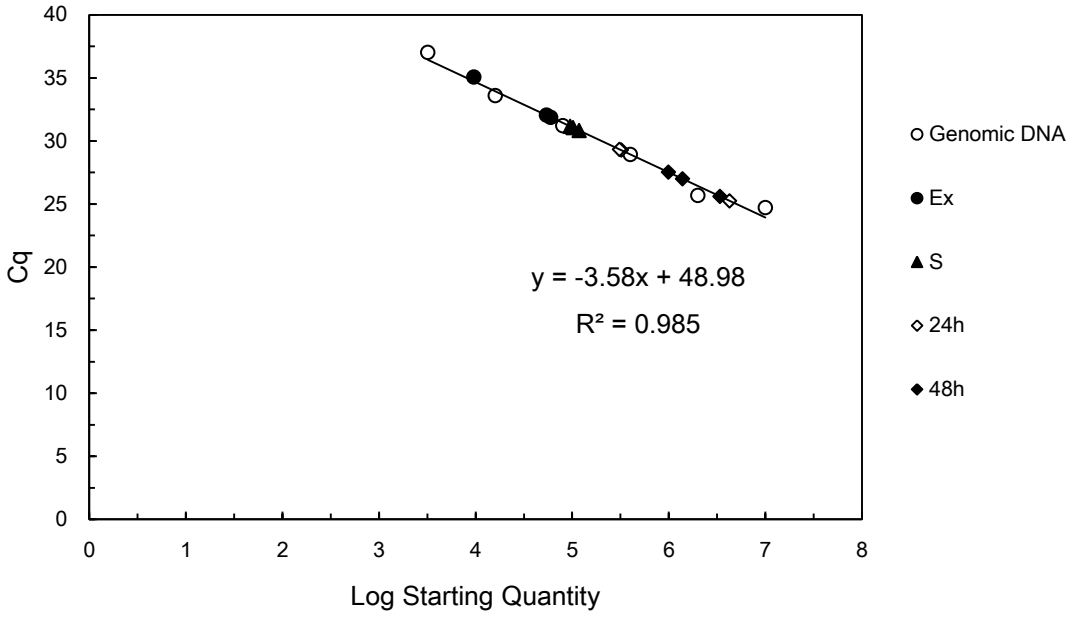


Figure 13. *tomB* Standard curve and Cq values of cDNA from exponential phase to 48h of incubation. White circle symbols indicate the Cq values of the serially diluted genomic DNA which were used as the templates to construct standard curve. The black cycles indicate the Cq values of cDNA in exponential phase. The black triangles indicate Cq values of cDNA in early stationary phase. The white diamonds indicate Cq values of cDNA at 24 h of incubation. The black diamonds indicate the Cq values of cDNA at 48 h of incubation. The amplification efficiency (E) was 90.3% and the coefficient of determination (R^2) was 0.985. The dynamic range was between 3.2 x 10^3 and 1.0 x 10^7 gene copies. The Cq values of tested cDNA samples were within the dynamic range.

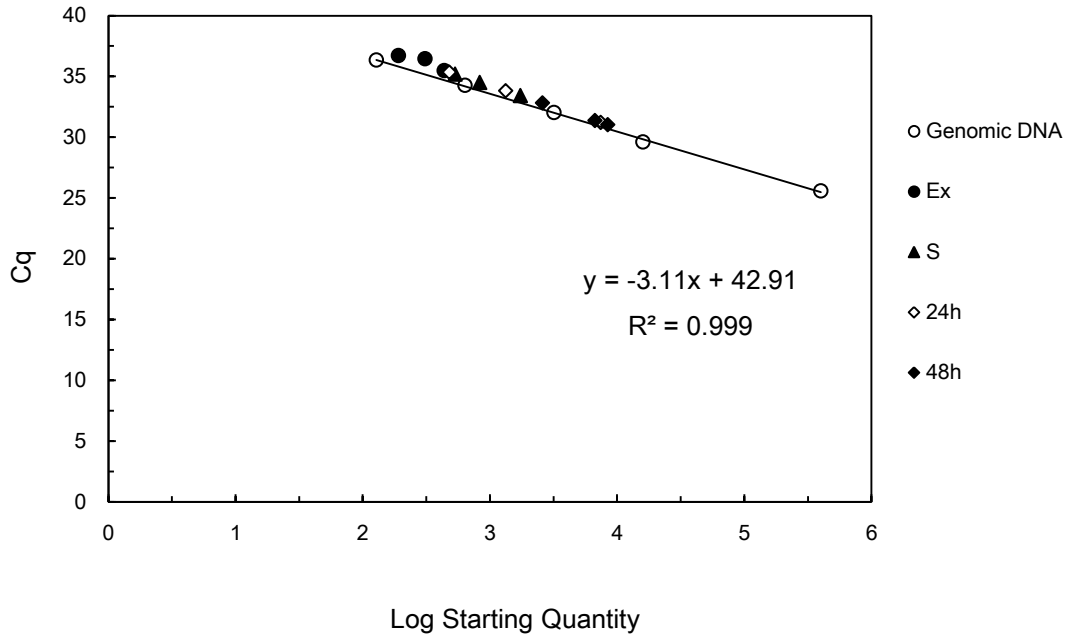


Figure 14. *emrK* Standard curve and Cq values of cDNA from exponential phase to 48h of incubation. White circle symbols indicate the Cq values of the serially diluted genomic DNA which were used as the templates to construct standard curve. The black cycles indicate the Cq values of cDNA in exponential phase. The black triangles indicate Cq values of cDNA in early stationary phase. The white diamonds indicate Cq values of cDNA at 24 h of incubation. The black diamonds indicate the Cq values of cDNA at 48 h of incubation. The amplification efficiency (E) was 110.0% and the coefficient of determination (R^2) was 0.999. The dynamic range was between 1.28×10^2 and 4.0×10^5 gene copies. The Cq values of tested cDNA samples were within the dynamic range.

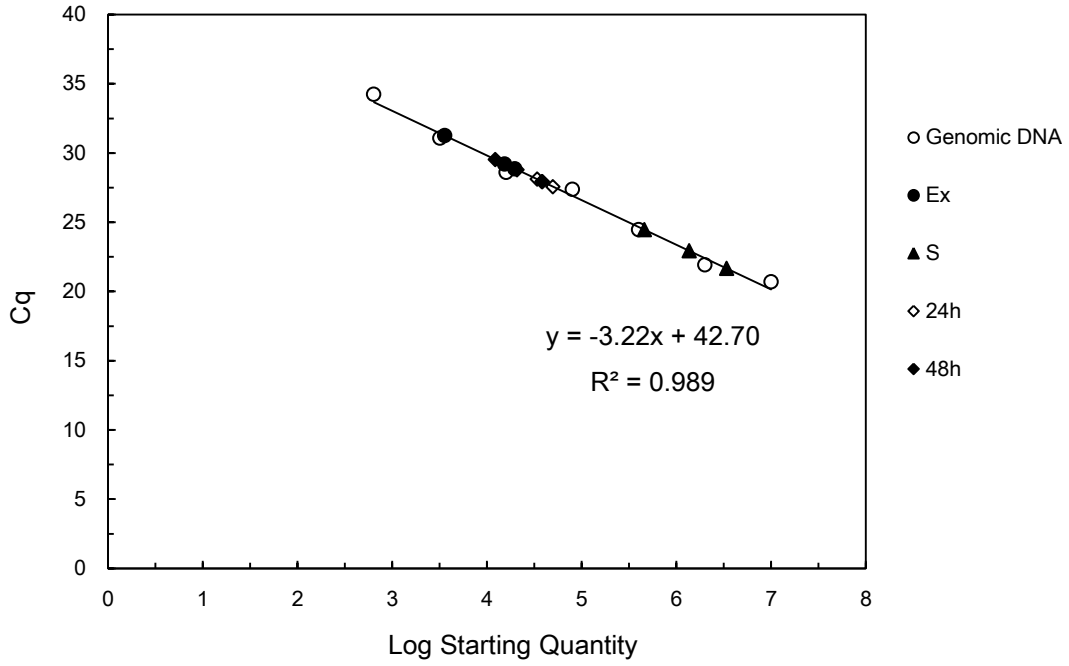


Figure 15. *osmY* Standard curve and Cq values of cDNA from exponential phase to 48h of incubation. White circle symbols indicate the Cq values of the serially diluted genomic DNA which were used as the templates to construct standard curve. The black cycles indicate the Cq values of cDNA in exponential phase. The black triangles indicate Cq values of cDNA in early stationary phase. The white diamonds indicate Cq values of cDNA at 24 h of incubation. The black diamonds indicate the Cq values of cDNA at 48 h of incubation. The amplification efficiency (E) was 104.4% and the coefficient of determination (R^2) was 0.989. The dynamic range was between 6.4 x 10^2 and 1.0 x 10^7 gene copies. The Cq values of tested cDNA samples were within the dynamic range.

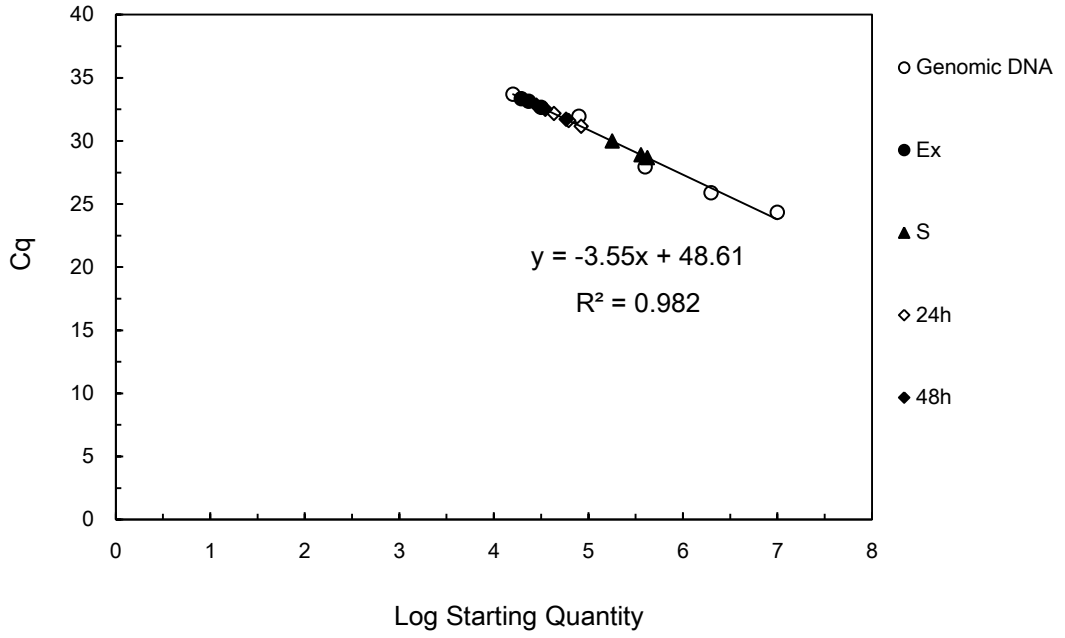


Figure 16. *bolA* Standard curve and Cq values of cDNA from exponential phase to 48h of incubation. White circle symbols indicate the Cq values of the serially diluted genomic DNA which were used as the templates to construct standard curve. The black cycles indicate the Cq values of cDNA in exponential phase. The black triangles indicate Cq values of cDNA in early stationary phase. The white diamonds indicate Cq values of cDNA at 24 h of incubation. The black diamonds indicate the Cq values of cDNA at 48 h of incubation. The amplification efficiency (E) was 91.4% and the coefficient of determination (R^2) was 0.982. The dynamic range was between 1.6 x 10^4 and 1.0 x 10^7 gene copies. The Cq values of tested cDNA samples were within the dynamic range.

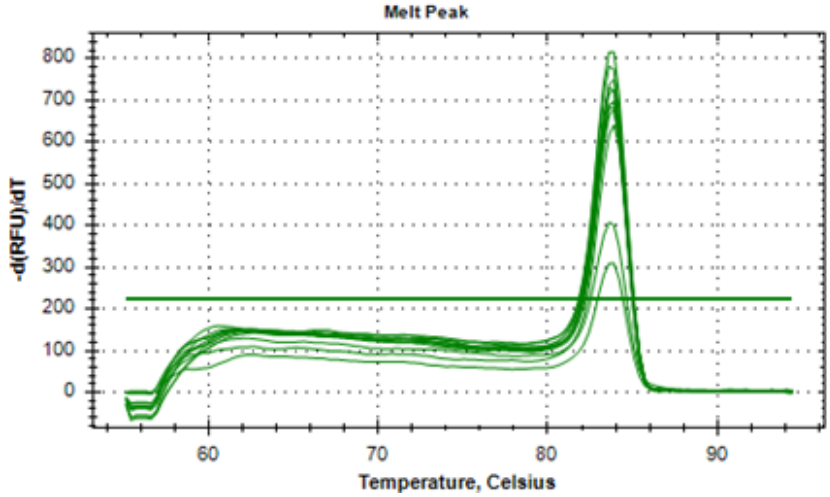


Figure 17. Melt curve of *rrsA* amplification. The single peak indicated the specificity of *rrsA* amplification product. This graph was exported from Bio-Rad CFX Manager™.

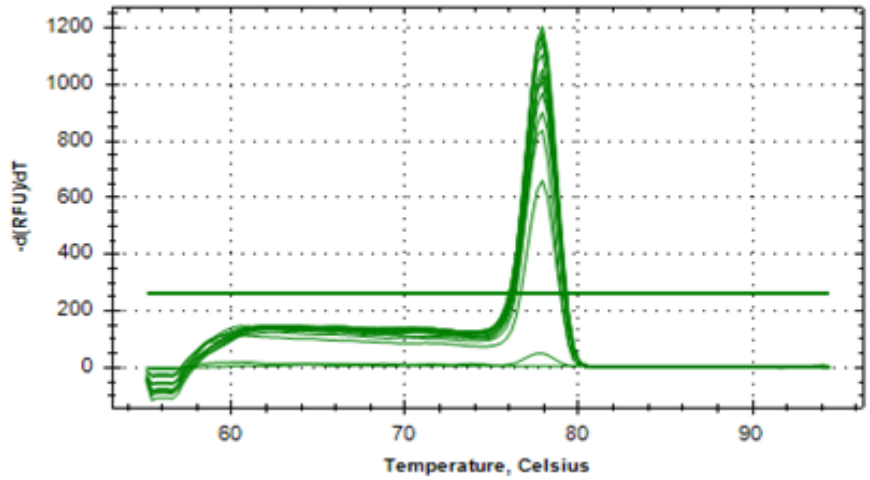


Figure 18. Melt curve of *hha* amplification. The single peak indicated the specificity of *hha* amplification product. This graph was exported from Bio-Rad CFX Manager™.

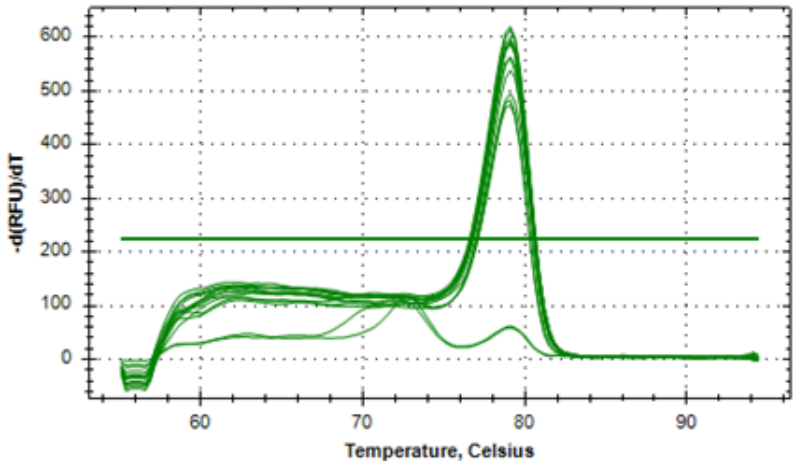


Figure 19. Melt curve of *tomB* amplification. The single peak indicated the specificity of *tomB* amplification product. This graph was exported from Bio-Rad CFX Manager™.

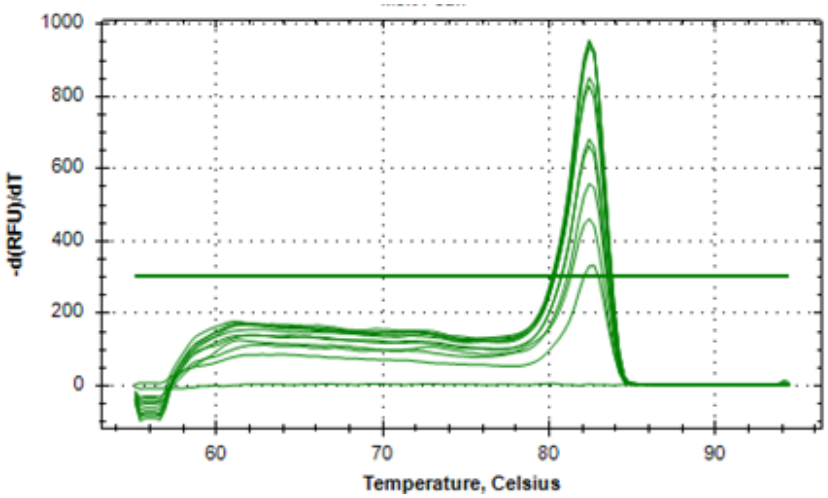


Figure 20. Melt curve of *emrK* amplification. The single peak indicated the specificity of *emrK* amplification product. This graph was exported from Bio-Rad CFX Manager™.

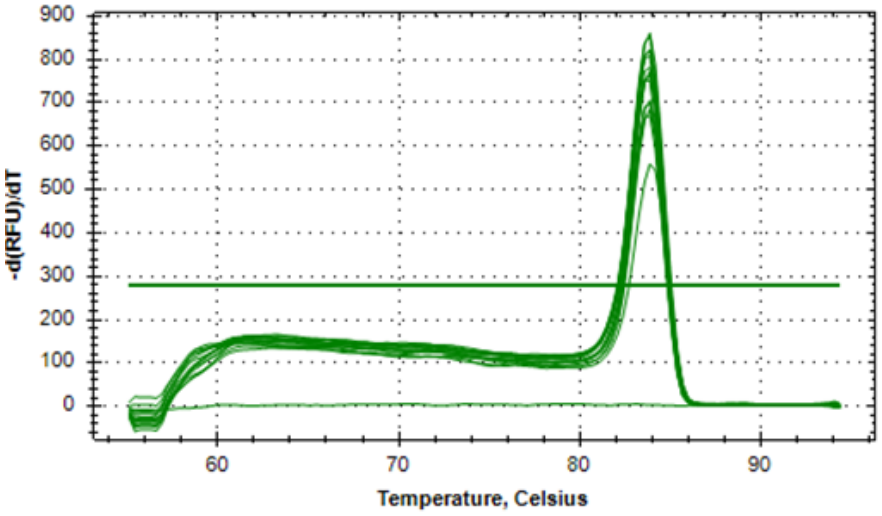


Figure 21. Melt curve of *osmY* amplification. The single peak indicated the specificity of *osmY* amplification product. This graph was exported from Bio-Rad CFX Manager™.

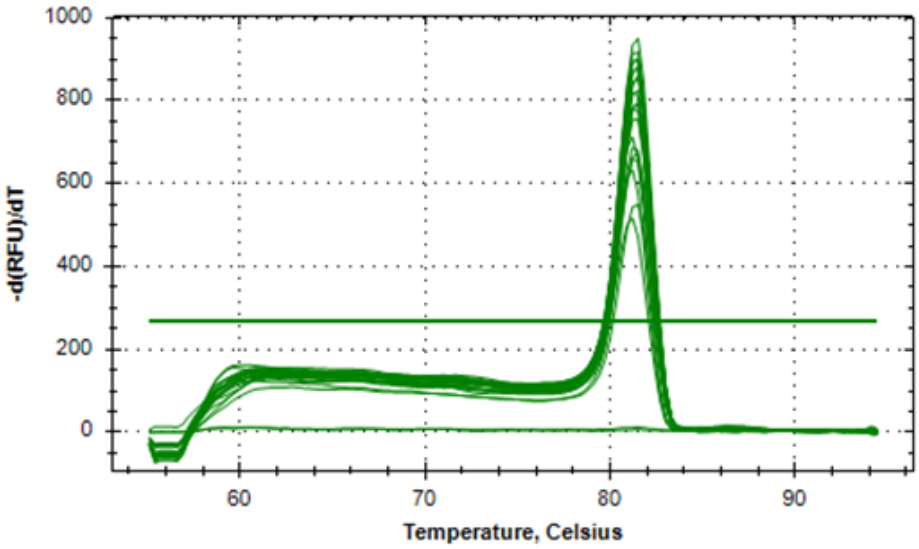


Figure 22. Melt curve of *bolA* amplification. The single peak indicated the specificity of *bolA* amplification product. This graph was exported from Bio-Rad CFX Manager™.

Appendix B. Standard operating protocol

Cell growth

1. Streak out from stocked *E. coli* MG1655 on a 1 x LB agar plate and incubate at 37˚C for overnight.
2. Inoculate a single colony into 1 x LB liquid culture and incubate at 37 ˚C with shaking at 200 rpm for overnight.
3. Subculture strain in a dilution at OD_{600} of 0.0001 into 50 ml 1 x LB, incubate at 37 ˚C with shaking at 200 rpm.
4. Take cells of OD_{600} of 1.0 from cultures at OD_{600} of 0.3, 1.5 and at 24 h and 48 h of incubation into 1.5 ml microcentrifuge tube.
5. Centrifuge at 12,000 x g for 5 min. Discard supernatant.
6. Flash-freeze in liquid nitrogen and store at -80 ˚C.

RNA isolation

(According to Norgen Biotek Total RNA Isolation protocol, #17200)

1. Add 1 mg/ml lysozyme-containing TE buffer directly into the frozen bacterial pellet and vortex for 10 s.
2. Incubate at room temperature for 5 min.
3. Add 300 µl of lysis solution and vortex for at least 10 s.
4. Add 200 µl of 100% ethanol to the lysate. Vortex for 10 s.
5. Assemble a column with the collection tubes.

6. Apply 600 µl of lysate with ethanol onto the column and centrifuge at 14,000 x g for 1 min.
7. Discard the flowthrough. Reassemble the spin column with its collection tube.
8. Apply 400 µl of wash solution to the column and centrifuge for 2 min.
9. Discard the flowthrough and reassemble the spin column with its collection tube.
10. Apply 100 µl of the RNase-free DNase I solution (20 U) to the column and centrifuge at 14,000 x g for 1 min.
11. Pipette the flowthrough that is present in the collection tube back onto the top of the column.
12. Incubate the column assembly at 37 ˚C for 20 min.
13. Apply 400 µl of wash solution to the column and centrifuge for 1 min.
14. Discard the flowthrough and reassemble the spin column with its collection tube.
15. Apply 400 µl of wash solution to the column and centrifuge for 1 min.
16. Discard the flowthrough and reassemble the spin column with its collection tube.
17. Spin the column for 2 min.
18. Place the column into a flesh 1.7 ml Elution tube.
19. Add 20 µl of Elution solution to the column.
20. Centrifuge at 2,000 x g for 2 min, followed by 1 min at 14,000 x g.
21. Add another 20 µl of Elution solution to the column.
22. Centrifuge at 2,000 x g for 2 min, followed by 1 min at 14,000x g.
23. Measure RNA and DNA concentration with Qubit fluorometer.

DNA isolation

(According to Bacterial genomic DNA isolation protocol form Norgen Biotek, #17900)

1. Transfer 1 ml of bacterial culture to 1.5 ml microcentrifuge tube and centrifuge at 14,000 x g for 30 s to pellet the cell. Discard supernatant
2. Add 250 µl of resuspension solution to the cell pellet. Resuspend the cells by gentle vortexing.
3. Add 250 µl of the lysis solution and 12 µl of proteinase K (20 ug/µl) to the cell suspension. Mix well by gentle vortexing and incubate at 55˚C for 30 min.
4. Add 500 µl of binding solution to the lysate and mix well by gentle vortexing.
5. Assemble a spin column with the collection tube. Apply 750 µl of the mixture to the spin column assembly. Centrifuge at 5,200 x g for 1 min.
6. Discard the flowthrough and apply the rest of the mixture to the column, centrifuge at 5,200 x g for 1 min.
7. Discard the flowthrough and reassemble the spin column with its collection tube.
8. Apply 500 µl of wash solution, centrifuge at 14,000 x g for 1 min. Discard the flowthrough and reassemble the spin column with its collection tube. Repeat this step once.
9. Spin the column for another 2 min.
10. Assemble the spin column with the 1.7 ml Elution tube.
11. Add 40 µl of Elution Buffer to the column. Centrifuge at 3,000 x g for 1 min. Centrifuge at 1,4000 x g for another 2 min.

cDNA synthesis

1. Add 4 µl of 5 X iScript reaction mix, 160 ng of RNA and nuclease free water to make up the volume to 20 µl.
2. Incubate the reaction mixture at 25 ˚C for 5 min, 42 ˚C for 30 min and 85 ˚C for 5 min.
3. Store cDNA at -20 ˚C.

qPCR reaction

1. Make 5-fold serial dilutions of genomic DNA starting from 43 ng.
2. Make qPCR mixture containing 8 ng of template, 4 µl of 2 X SsoFast^{TM} EvaGreen Supermix (Bio-Rad), 0.2 µM of forward and reverse primer and 3 µl of H_2O.
3. Run qPCR in the program: enzyme activation at 95 ℃ for 2 min, followed by 40 cycles of denaturation at 95 ℃ for 5 s and annealing/extension at 62 ℃ for 3 s. A melt curve assay was then performed with increments of 0.3 ℃ ranged from 55-95 ℃ for 5 s.

Gel analysis

1. Aliquot 3 µl of samples into 3 µl of 1 x loading dye and load into 0.8% agarose gel containing 0.1 ng/ml ethidium bromide.
2. Run the gel at 120 V for 40 min.
3. Visualize RNA or PCR products using a gel documentation system.

RNA isolation (modified small-scale phenol-chloroform)

1. Thaw bacteria pellet on ice.

2. Add 0.5 ml of boiling 2% SDS/16 mM EDTA (heated in microwave). Thoroughly mix the bacterial pellet by inversion.
3. Add 1 ml of 65˚C acid phenol. Mix well by inversion.
4. Place tubes in 65˚C water bath for 3 min, inverting every 30 s.
5. Place the tubes on ice for 10 min, inverting occasionally.
6. Centrifuge tubes at 13,000 x g for 10 min.
7. Transfer aqueous phase to 1.5 ml microcentrifuge tube. Add 80 U DNase I and incubate at 37 ˚C for 20 min.
8. Add DEPC water to make up volume of the aqueous phase (to 0.5 ml).
9. Add 1 ml acid phenol. Mix well by inversion several times.
10. Centrifuge at 13,000 x g for 10 min, transfer aqueous phase to 1.5 ml microcentrifuge tube.
11. Repeat step 9 and 10 once or until no white mater is visible at the interface.
12. Add 1 ml of room temperature phenol-chloroform-isoamyl (25:24:1). Mix well by inversion.
13. Centrifuge tubes at 13,000 x g for 5 min and transfer aqueous phase to 1.5 ml microcentrifuge tube.
14. Add 1 ml of room temperature chloroform-isoamyl alcohol (24:1). Mix well by inversion.
15. Centrifuge tubes at 13,000 x g for 5 min and transfer aqueous phase to 1.5 ml microcentrifuge tube.
16. Add 40 µl of 3 M sodium acetate (pH 5.2). Mix well by inversion.

17. Add 1 ml of absolute ethanol. Mix well by inversion.
18. Incubate at -20 ˚C for 60 min or overnight.
19. Centrifuge at 13,000 x g for 30 min at 4 ˚C. Discard supernatant.
20. Wash pellet with 0.5 ml of 70% ethanol. Centrifuge at 13,000 x g for 15 min at 4 ˚C. Discard supernatant.
21. Dry RNA pellet for 15 min at 37 ˚C.
22. Add 20 µl of DEPC water to Resuspend RNA pellet.
23. Store at -20 ˚C.
24. Measure RNA and DNA concentration with Qubit fluorometer.

DNase treatment

1. Add 10 µl of DNase 10 x buffer, 20 µl of RNA sample and 5µl of DNase I (1 U/µl).
2. Tap to mix, incubate at 37 ˚C for 30 min.
3. Add 500 µl of DEPC water and 0.1 volume of sodium acetate (3 M).
4. Extract with equal volume of hot phenol-chloroform 3 times, each time add water to make up the loss of volume.
5. Precipitate RNA with 2.5 volume of 100% ethanol and 0.1 volume of sodium acetate (3 M).
6. Vortex for 15 s.
7. Centrifuge at 12,000 g for 15 min.
8. Add 700 µl of 75% ethanol, vortex 15 s.
9. Incubate at -20 ˚C for 5 min.
10. Centrifuge at 12,000 g for 10 min.

11. Dry pellet in the air for 15 min.

12. Resuspend RNA in 20 µl of DEPC water and store at -20 ˚C.

Equipment:

Centrifuge (Labnet)

Water bath (VMR Co.)

Incubator (Precision scientific Co.)

Qubit fluorometer (Invitrogen)

qPCR (CFX96 Touch™ Real-Time PCR Reaction System (Bio-Rad)

Appendix C. Thesis defense

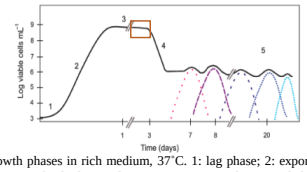
Identification of Preferentially Expressed Genes in *Escherichia coli* in Late Stationary Phase

Lingzi Kong

Sept. 20, 2013

Background: life cycle

- *E. coli* is a common intestinal bacterium and can survive in nature for long periods
- *E. coli* life cycle consists of 5 stages in batch culture



Bacterial growth phases in rich medium, 37˚C. 1: lag phase; 2: exponential phase; 3: stationary phase; 4: death phase; 5 long-term stationary phase. (Finkel, 2006)

E. coli cells in stationary phase

- Morphological changes
  - Small and spherical
  - Cell envelope
  - Condensed nucleoid
  - RNA
    - 20%~40% total RNA is lost
    - rRNA degradation
    - Ribosomal proteins remain stable
    - Induced by limited nutrients

E. coli cells in stationary phase

- Physiological changes
  - Important regulators
    - ppGpp, RpoS
  - Metabolic activities
    - Decreased protein synthesis rate
      - Dimerization of ribosome
    - Synthesis of secondary metabolites
      - Antibiotics, QS (quorum sensing) molecules
  - Programmed cell death
    - Provide nutrient and limit the spread of viral infection
  - Resistance to external stress
  - High genomic mutation rate

Objective

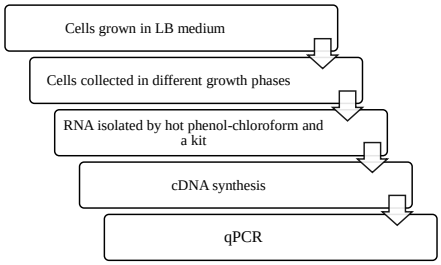
- *E. coli* cells mediate gene expression to alter the morphology and physiology
- Identify the preferentially expressed genes in *E. coli* in late stationary phase
  - General profile of gene expression determined by previous microarray
  - Validation of microarray data by qPCR
    - Target genes
  - RNA-seq

Target genes

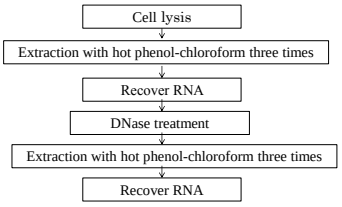
Gene	Function	Fold change			
		Exponential	Stationary	24 h	48 h
hha	High hemolysin activity modulator	1.0	0.3	1.0	7.8
tomB	Attenuate the toxicity of Hha	1.0	1.0	5.0	24.0
emrK	Multidrug efflux pump	1.0	1.4	2.7	13.1
osmY	Osmotic stress regulation	1.0	20.0	345.0	52.0
bolA	Alter morphology	1.0	14.0	52.0	8.0

The fold change was determined by dividing the expression intensity of target gene in each growth phase with that in exponential phase. The microarray data of gene expression in exponential phase and early stationary phase was obtained from T. Dong [14]. The microarray data of gene expression in 1 day and 2 days was obtained from S. Chiang.

Workflow

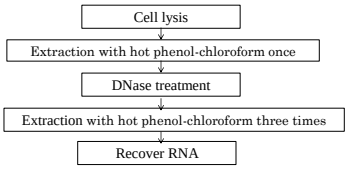


Traditional hot phenol-chloroform method



- Low RNA recovery in late stationary phase

Modified hot phenol-chloroform method



- Low RNA recovery
  - Yield RNA was calculated: 10 pg (RNA amount per cell) x cell numbers (OD_{600} of 1.0 = 8 x 10^8 cells/ml)
  - 24 h: 3.65% RNA recovery; 48 h: 2.50% RNA recovery

DNase titration

Cells collected in OD_{600} of 1.0 at 24 h of incubation

DNase amount (U)	RNA concentration (ng/µl)	RNA recovery (%) [a]	DNA concentration (ng/µl)	Genomic DNA contamination (%) [b]
0	460	22	78	17
10	540	22	24	4
20	576	23	14	3
40	448	18	12	3
60	456	18	12	3
80	373	14	10	3

[a] RNA recovery (%): recovered RNA amount / 80 µg
[b] Genomic DNA contamination: DNA concentration / RNA concentration

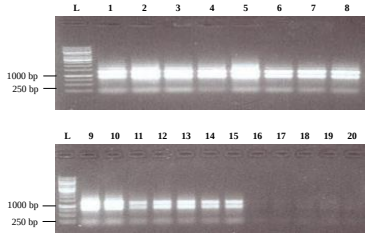
RNA extraction: kit

- RNA extracted from cells of OD_{600} of 1.0
  - Yield RNA amount was 80 µg
- 20 U of DNase was used to remove genomic DNA on column

Time	RNA concentration (ng/µl)	RNA amount (µg)	RNA recovery (%) [a]	Genomic DNA contamination (%) [b]
5 h	1203.6 ± 56.5	47.8 ± 2.3	58.3 ± 4.1	2.0 ± 0.0
7 h	932.4 ± 22.1	37.0 ± 1.1	48.5 ± 1.9	2.0 ± 0.0
24 h	303.6 ± 11.5	12.2 ± 0.5	27.3 ± 2.0	2.4 ± 0.4
48 h	30.8 ± 1.4	1.0 ± 0.0	1.6 ± 0.1	9.2 ± 0.4

[a] RNA recovery (%): recovered RNA amount / 80 µg
[b] Genomic DNA contamination: DNA concentration / RNA concentration

RNA integrity test on gel



Lane 1-5: RNA extracted from exponential cultures
Lane 6-10: RNA extracted from early stationary phase cultures
Lane 11-15: RNA extracted from 24 h cultures
Lane 16-20: RNA extracted from 48 h cultures
Lane L: 1 kb ladder

RNA integrity check on 0.8% agarose gel.
RNA was extracted from *E. coli* MG1655 from exponential phase to 48 h of incubation in 5 biological replicates by using a commercial kit.

Reasons for the low RNA recovery

- RNA loss during isolation
  - Comparison between the amount of yield and recovered RNA in exponential phase
- Yield RNA amount ≠ 80 µg
  - Due to RNA degradation and cell death
  - RNA amount was determined by measuring RNA concentration after cell lysis
  - Comparison between the yield RNA amount in exponential phase and 48 h of incubation

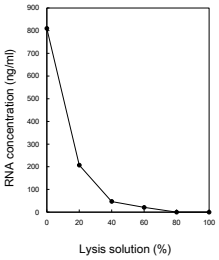
PLYR concentration measurement

- RNA concentration was measured with the Qubit fluorometer

Time	Growth phase	PLYR concentration (ng/µl)	PLYR amount (ng)	Recovered RNA[b] concentration (ng/µl)	Recovered RNA amount (ng)
5 h	Exponential	6.62	2648	1280	51200
48 h	2 days	0.33	132	37	1488

[a] PLYR (post lysis yield RNA: RNA in the cells after lysis
[b] Recovered RNA: extracted RNA after whole isolation was finished

Lysis solution effect on the measurement of yield RNA



Lysis solution effect on the measurement of RNA concentration.

60% lysis solution results in 40-fold RNA concentration underestimation.

RNA recovery

- During isolation, RNA was resuspended in 600 µl of 75% lysis solution
- At least 40-fold RNA was underestimated

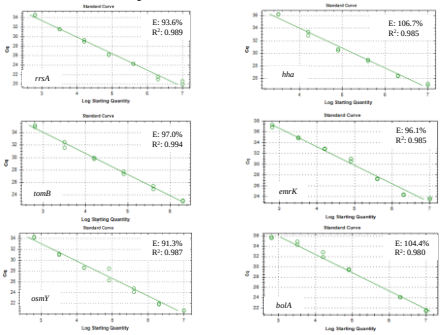
Time	Estimated PLYR amount (ng) [a]	Actual PLYR amount (ng) [b]	Recovered RNA amount (ng)	RNA recovery (%)
5 h	2648	105000	51200	48.8
48 h	132	5000	1488	28.4

[a] Estimated PLYR amount = RNA concentration (Qubit fluorometer) x 600 µl
[b] Actual PLYR amount = Estimated PLYR amount x 40

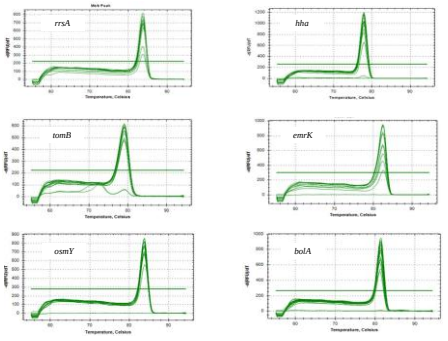
Primer assay

- Amplification efficiency (E) test
  - Standard curve
  - 90 – 110% is required
  - R^2 (coefficient of determination) > 98%
- Melt curve
  - Identify the specificity of amplified product

Standard curves of primers



Melt curves of primers



Quantification of gene expression by qPCR

- Relative standard curve method
- *rrsA* with constant expression in late stationary phase was used to normalize target gene expression
- Fold change was used to represent the expression change between each growth phase and exponential phase

Quantification of *tomB* expression from exponential phase to 48 h of incubation

Growth phase	Expression (log_2 transformed)		Normalized expression value [a]	Fold change [b]	
	tomB	*rrsA*		qPCR	Microarray
Exponential	14.9 ± 0.9	22.5 ± 0.3	0.006	1	1
Early stationary	16.7 ± 0.1	22.7 ± 0.3	0.015	2	1
1 day	19.5 ± 1.3	21.1 ± 0.1	0.725	118	5
2 days	20.7 ± 0.5	19.1 ± 0.3	3.180	519	24

[a] The normalized expression of *tomB* in each sample was calculated using the expression of *tomB* divided by the expression of *rrsA*.
[b] Fold change was determined by using the normalized expression of *tomB* at each time point divided by the normalized expression of *tomB* at exponential phase.

Quantification of *bolA* expression from exponential phase to 48 h of incubation

- RpoS-dependent gene

Growth phase	Expression (log_2 transformed)		Normalized expression value [a]	Fold change [b]	
	bolA	*rrsA*		qPCR	Microarray
Exponential	14.6 ± 0.2	22.1 ± 0.4	0.005	1	1
Early stationary	18.2 ± 0.4	21.9 ± 0.1	0.083	16	20
1 day	15.9 ± 0.3	20.7 ± 0.2	0.037	7	345
2 days	15.2 ± 0.3	18.6 ± 0.4	0.097	19	3

[a] The normalized expression of *bolA* in each sample was calculated using the expression of *bolA* divided by the expression of *rrsA*.
[b] Fold change was determined by using the normalized expression of *bolA* at each time point divided by the normalized expression of *bolA* at exponential phase.

Validation of microarray data

Growth phase	*hha*	*tomB*	*emrK*	*osmY*	*bolA*
Exponential – early stationary	N	N	Y	Y	Y
Early stationary – 1 day	N	Y	Y	N	N
1 day – 2 days	Y	Y	Y	N	N

Y : Confirmed result
N : Discordant result

Reasons for the discordant results

- Less than 2-fold change determined by microarray
- Genomic DNA contamination and low abundance transcripts
- Lack of biological replicate in microarray experiment
- Inherent disadvantages
  - qPCR: low amplification efficiency in later cycles
  - Microarray: cross-hybridization between probes and target genes

Conclusions

- RNA recovery decreased over time
- The expression of *hha* increased from 24 h to 48 h of incubation.
- The expression of *tomB* increased from stationary phase to 48 h of incubation.
- The expression of *emrK* increased from exponential phase to 48 h of incubation.

Acknowledgement

- Dr. Schellhorn
- Dr. Gupta
- Dr. Kolasa
- All lab members, past and present
- My parents and friends

Thanks