GENE EXPRESSION IN *ESCHERICHIA COLI* DURING PROLONGED-INCUBATION

GENE EXPRESSION IN *ESCHERICHIA COLI* DURING PROLONGED-INCUBATION

BY BANSRI PATEL, M. Sc.

A Thesis Submitted to the School of Graduate Studies in Partial fulfilment of the Requirements for the Degree Master of Science

McMaster University

MASTER OF SCIENCE (2018) McMaster University

(Biology Department) Hamilton, Ontario

TITLE: Gene expression in *Escherichia coli* during prolonged-incubation

AUTHOR: Bansri Patel, M. Sc. (Sardar Patel University, India)

SUPERVISOR: Dr. Herb E. Schellhorn

NUMBER OF PAGES: viii, 165

ABSTRACT

Environments, where growth is limited by the availability of nutrients are common, for example, soil, water, or even host environments such as macrophages, can lack essential nutrients to support growth. As such, many bacteria spend most of their time in states of little or no growth due to starvation. The starved and growth attenuated state is now widely considered as an important physiological condition in bacterial pathogenesis and survival. Experiments studying stationary phase and adaptation mechanisms to non-optimal conditions lead to the discovery of RpoS as a growth phase-dependent sigma factor. Though RpoS controls many genes in the early stationary phase, it is not known whether RpoS is necessary for prolonged slow growth or not. In a previous study to identify genes controlled by RpoS, we found that a large fraction of the *E. coli* genome continues to increase in expression during prolonged starvation that does not require RpoS. This suggests that other growth-phase-dependent regulatory mechanism, in addition, to RpoS, may control prolonged stationary phase gene expression. In this study, we examined the abundance of transcripts to identify and characterize the genes that are preferentially expressed during prolonged-incubation phase. RpoS independent genes that are expressed in higher abundance during prolonged-incubation include iron acquisition genes, enterobactin biosynthesis, arginine degradation, and 2-methycitrate pathways enzyme coding-genes. Putative fimbriae genes associated with adhesion to biotic and abiotic surfaces are expressed as RpoS-dependent genes. Furthermore, several biofilm formation genes are expressed in planktonic cultures. Altogether, other regulators, in addition to RpoS, regulate the gene expression during the prolonged-incubation phase and the genes are likely to be important for survival during prolonged-incubation phase.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Escherichia coli is a model organism that provides insights into genetic and regulatory mechanisms that are shared and retained by all organisms (Blount, 2015, Penders *et al.*, 2006). Moreover, *E. coli* is critical for studies in biological engineering and industrial microbiology (Lee, 1996). *E. coli* exist in its primary and secondary habitats. The intestine of warm blooded animals is considered to be a primary habitat of *E. coli* which has an anaerobic environment that supports constant growth due to the steady influx of nutrients. In a secondary habitat such as soil, water, or sediments, *E. coli* faces fluctuations in environmental parameters such as pH, temperature and osmotic pressure (Winfield & Groisman, 2003). *E. coli* has the ability to survive under stressful conditions found in both primary and secondary habitats until *E. coli* encounters optimal conditions for growth (FLIN, 1987, Huai-Shu XU & ColwelP, 1982).

E. coli under stressful conditions not only differ at the transcriptome and proteome levels but also exhibit morphological diversity. For instance, exponential phase cells are motile and flagellated, while stationary phase cells lack flagella (Makinoshima *et al.*, 2003). *E. coli* expresses many survival and stress tolerance genes during distinct physiological growth stages such as biofilms (Schembri *et al.*, 2003, Ren *et al.*, 2004), persister cells (Shah *et al.*, 2006) and stationary phase (Patten *et al.*, 2004, Schellhorn HE, 1998). The ability of *E. coli* to adapt to stressful conditions allow it to colonize many environments, such as medical devices (Donlan & Costerton, 2002) and contributes to causing infections in humans (Ren *et al.*, 2004). The understanding of gene expression during distinct physiological growth stages of *E. coli* has greatly improved. However, our molecular knowledge of the entire starvation response during the stationary phase is still far from complete. The following is a discussion of the stationary phase planktonic cultures with a focus on the expression of genes require to survive under this condition.

1.1 Stationary phase of *E. coli*

E. coli possess regulatory networks that mediate cellular responses to stressful conditions (Hengge-Aronis, 1991b). The response can be specific for a given stress condition. For instance, the induction of the heat shock sigma factor, RpoH in response to a temperature shift from 30° C to 42° C (Bukau, 1993) and the up-regulation of genes responsible for stabilizing and refolding denatured proteins, as well as those genes required for proteolysis of misfolded proteins (Nonaka *et al.*, 2006). In contrast to the specific response, the stationary phase sigma factor RpoS is induced in response to multiple non-optimal conditions and controls about 10% of *E. coli* genome during entry into stationary phase (Weber *et al.*, 2005). Stress stimuli includes osmotic shock (Cheung *et al.*, 2003), oxidative stress (Schellhorn, 1995), heat stress (Hengge-Aronis *et al.*, 1991) and entry into stationary phase (Hengge-Aronis, 1991b, Patten *et al.*, 2004).

1.1.1 RpoS mediates stationary phase adaptation

Microarrays (Lacour & Landini, 2004, Patten *et al.*, 2004, Weber *et al.*, 2005) and transcriptional reporters (Schellhorn *et al.*, 1998) have been used to examine the expression levels of the genes within the RpoS regulon. Genes responsible for protein synthesis such as the ribosome-associated protein RpsV (*sra*) and the initiation factors IF-1 (*infA*) are induced by RpoS. Moreover, the RpoS also controls indole production, acting as an extracellular signaling molecule through regulation of *tnaA* gene encoding tryptophanase enzyme (which converts tryptophan to indole) (Lacour & Landini, 2004). RpoS also controls the transcription of several genes responsible for biosynthesis of the signal molecule c-di-GMP, thus promoting production of adhesion and cell aggregation factors while reducing flagella-mediated cell motility (Sommerfeldt *et al.*, 2009).Consistently, RpoS negatively regulates flagellar genes for motility, genes encoding for some enzymes of the TCA cycle, genes for Fe-S clusters proteins, and *rac* prophage genes (Patten

et al., 2004). RpoS dependent regulation has been examined under three different conditions: the transition to stationary phase in LB, 20 minutes after addition of NaCl (0.3 M) in minimal medium to cause an osmotic upshift, and 40 minutes after acidification of rich media by MES. A total of 140 genes are positively regulated under all three conditions. These genes comprised the core set of genes regulated by RpoS (Weber *et al.*, 2005). Data suggests that the overall set of genes controlled by RpoS depends on growth conditions and is also affected by additional regulators (Weber *et al.*, 2005).

Expression of the glutamate-dependent decarboxylases *gadA* and *gadB,* is RpoS-dependent in stationary phase, but not in the exponential phase during acid stress. Moreover, the cAMP-CRP protein also directly controls several RpoS regulated genes. The expression profiles under diverse nutrient levels are different (Dong & Schellhorn, 2009a) may be due to the constitutively high level of RpoS in minimal media or additional regulatory proteins that interact with RpoS. The overlap between the complex regulatory network of RpoS regulated genes and other global regulatory regulons such as cAMP and Lrp protein support the latter possibility (Weber *et al.*, 2005). Furthermore, the composition of the RpoS regulon differs among the strains of *E. coli.* For instance, in the exponential phase of pathogenic *E. coli* strain, the expression profile of about 1000 genes is affected by RpoS deletion, while only 11 genes are affected in the laboratory strain (Dong & Schellhorn, 2009b). Moreover, the gene expression level of *tnaA* (tryptophanase) is positively regulated by RpoS, while, *tnaA* is negatively regulated by RpoS in laboratory strains during stationary phase (Patten *et al.*, 2004). The TCA cycle genes and the motility genes are negatively regulated by RpoS in laboratory strain, but no differential expression is observed in pathogenic strain. This suggests that even the negative regulation by RpoS is strain specific. In pathogenic *E.*

coli, 80% of genes expressed are strain specific implies that the composition of the RpoS regulon is highly divergent (Dong & Schellhorn, 2009a).

1.1.2 6S RNA and Rsd, regulators of stationary phase

In stationary phase, the concentration of RpoS is lower than the housekeeping sigma factor RpoD and RpoS also has lower affinity to RNAP than RpoD (Ishihama, 2000). As such, maximal induction of the RpoS regulon is achieved by regulated inhibition of RpoD activity combined with an increase in cellular levels of RpoS. Two regulators, the 6S RNA and Rsd, are growth phase-dependent but RpoS-independent these inhibit RpoD activity in stationary phase. There are several genes expressed during the stationary-phase that do not require RpoS for their expression (Hengge-Aronis, 1991a, Christensen-Dalsgaard *et al.*, 2010a, Kim & Wood, 2010), however, the genes mention in Table 1 are identified in different studies.

The *rsd* gene encodes a stationary phase specific anti-sigma factor that binds RpoD (Ishihama, 1998) this results in a concomitant increase in RpoS-dependent promoters (Mitchell *et al.*, 2007). Interestingly, Rsd levels are high even in exponential phase of growth however inhibitory effects of Rsd are only seen in stationary phase (Piper *et al.*, 2009). 6S RNA is an abundant RNA regulator encoded by *ssrA* gene, that resembles a promoter sequence and therefore causes sequestration of RpoD bound RNAP downregulating gene expression at a global scale during transition to stationary phase (Barrick *et al.*, 2005). The levels of 6S RNA increase from about 1000 copies per cell to 10,000 copies by 24h of incubation with a constant increase. Furthermore, this regulation is not affected by an *rpoS*::Tn10 mutation. 6S RNA specifically binds RpoD-RNAP and inhibits its activity without affecting RpoS-RNAP activity (Wassarman & Storz, 2000). Despite the abundance of 6S and given that majority of RNAP is bound by 6S in stationary phase, not all promoters are affected. A weak -35 element determines promoter specificity at 6S

RNA dependent promoters (Cavanagh *et al.*, 2008). Moreover, in exponential phase *guaD-ygfQ* operon (a guanine deaminase and a transporter) and *tdcABCDEFG* operon (encoding serine and threonine degradation proteins) are highly repressed by 6S RNA. Interestingly, stress response proteins Dps, UspF, UspG are downregulated in exponential phase. During early stationary phase, interacting H-NS related proteins (Hha, YdgT, and SlyA), the tryptophan transporter (Mtr) are downregulated by 6S RNA while the genes related to translation are upregulated (Neusser *et al.*, 2010). Furthermore, stress-related proteins are also down-regulated during exponential phase (Thomas Neusser, 2010). Interestingly, 6S RNA mutants showed reduced expression of RpoS dependent genes in exponential phase, mid-exponential phase, and stationary phase but not in late stationary phase. This is not due to a reduced level of RpoS protein but may be due to reduced activity of RpoS in late stationary phase (Lal *et al.*, 2016).

1.1.3 The stringent response

The stringent response is a bacterial stress response that controls adaptation to nutrient deprivation and is activated by several different starvation and stress signals. The molecular hallmark of this response is the synthesis of the small molecule called guanosine 5′, 3′ bispyrophosphate (ppGpp). Two gene *relA* and *spoT* activates the synthesis of ppGpp under starvation condition. The ppGpp signalling molecule controls replication, transcription, and the activity of the enzymes of the stress response (Boutte & Crosson, 2013). Upon transition from the exponential to the stationary phase of growth, a sharp drop in rRNA synthesis is observed. No role of RpoS is observed during this down-regulation of protein synthesis (Michal Aviv, 1996). In stationary phase, about 40% of 70S ribosomes are converted into non-active 100S dimers, by the ribosome modulation factor that is encoded by the *rmf* gene (Kirawada, 1990). The expression of the *rmf* gene is RpoS independent (Masahiro Yamagishi, 1993), but it requires ppGpp. Rmf

proteins inactivate excess ribosomes and promote viability in stationary phase during non-optimal conditions, such as heat and osmotic shock (Niven & El-Sharoud, 2008). The alternate sigma factor E (σ^{24}/RpoE) gene is induced in response to extracytoplasmic stress and its activity increases in a growth phase-dependent manner. The expression of RpoE depends on ppGpp but is independent of RpoS (Costanzo & Ades, 2006). Sigma factor E induces the expression of genes for phase-specific cell lysis and controls transcription of other genes, including genes for cell envelope formation, cellular processes, and hypothetical protein-coding genes (Kabir *et al.*, 2005). During a prolonged starvation condition, the expression of *sspA* (stringent starvation protein A) increases and affects the protein synthesis and growth. SspA expression also increases during glucose, nitrogen, phosphate and amino acid starvation (Williams *et al.*, 1994). The promoter of the *sspA* gene is similar to gearbox promoters, and the expression of this gene requires ppGpp (Williams *et al.*, 1994) and not RpoS (Patten *et al.*, 2004). SspA inhibits the accumulation of the global regulator H-NS during stationary phase and play an essential for cell survival during acid-induced stress (Hansen *et al.*, 2005). Expression of universal stress genes *uspA, uspC, uspD* and *uspE* is RpoS-independent, but requires ppGpp. UspA is a general stress response gene induced in condition that elicit a reduction in growth rate. UspC, UspD, and UspE are paralogs of UspA that play non-redundant roles and are regulated similarly: all of the proteins are induced during glucose, phosphate, and nitrogen limitation as well as during treatment with mitocyin C. Single deletion mutants in these genes have reduced viability when exposed to UV radiation (Gustavsson, 2002).

Table 1: RpoS-independent genes expressed during stationary phase.

Genes	Functions	References
rmf (Ribosome modulation factor)	Converts about 40% 70S to 100S dimers	(Kirawada, 1990) (Masahiro Yamagishi, 1993)
rsd	Anti-sigma factor	(Piper *et al.*, 2009) (Ishihama, 1998)
uspA, *uspC*, *uspD*, *uspE* (Universal stress proteins)	*uspD, uspE, uspC* have non-redundant roles in response to glucose, phosphate, nitrogen starvation	(Gustavsson, 2002)
sspA (Stringent starvation protein A)	Affects gene expression during extended incubation	(Williams *et al.*, 1994) (Hansen *et al.*, 2005)
mqsR	Toxin mRNA interferase that promotes persistence and biofilm formation	(Shah *et al.*, 2006) (Christensen-Dalsgaard *et al.*, 2010a)
cspD	Growth phase-dependent central regulator in persister cells.	(Kim & Wood, 2010)
ssrS (6S RNA)	Growth phase-dependent regulatory RNA that down regulates the RpoD dependent genes in stationary phase	(Hofmann *et al.*, 2011)
rpoE (sigma factor E)	Induced phase-specific cell lysis	(Costanzo & Ades, 2006) (Kabir *et al.*, 2005)
mcbA	Synthesize colonic acid	(Hengge-Aronis, 1991a) (Zhang *et al.*, 2008)

1.1.4 Persister cells in stationary phase

Persister cells reach about 1% of the total cell number in stationary phase planktonic cultures and in biofilms (Vega *et al.*, 2012). The up-regulation of the Toxin-antitoxin system (TA system) is the characteristic feature of persister cells. Persister cells form a subset of the dormant and non-growing phenotypic variants of the general cell population that possess a low enough metabolism to survive antibiotic treatment (Wood *et al.*, 2013). The transcriptome profile of persister cells resembles the exponential phase of planktonic cultures, but the genes for energy production and flagellar expression are repressed. Stationary phase-specific catalase *katE* and

other stationary phase-specific genes, such as *bolA* and *osmY* are highly repressed in persister cells (Shah *et al.*, 2006). TA modules such as *yafQ-dinJ* and *yoeB-yefM* are highly up-regulated in persister cells, and also expressed in stationary phase cells. Toxin MqsR is the most induced protein in persister cells. The overproduction of *mqsR* gene leads to growth arrest and increases microbial resistance. The tolerance of the overproduced *mqsR* strain to the antibiotic is similar to persisters cells tolerances (Shah *et al.*, 2006). The overproduction of *mqsR* gene does not lyse cells, but rather causes a reversible inhibition of growth that is overcome by MqsA corresponding anti-toxin protein (Kasari *et al.*, 2010). Upon glucose and amino acid starvation, expression of the *mqsR* gene is RpoS-independent (Christensen-Dalsgaard *et al.*, 2010a).

Upon glucose starvation in the stationary phase, there is an expression of the CspD toxin (homology to cold shock protein CspA). The transcription of *cspD* gene is regulated by ppGpp, which is independent of RpoS. CspD inhibits DNA replication and acts as an RNA/DNA chaperone at physiological temperatures (Inouye, 1997). MqsR toxin is an activator of *cspD* gene and subsequently increased persister cell formation. However, its corresponding antitoxin MqsA represses the expression of *cspD* gene (Kim *et al.*, 2010). Additionally, the antitoxin MqsA represses RpoS-mediated general stress response and c-di-GMP signalling, responsible for promoting adhesion. MqsA, therefore promotes motility. Repression of the general stress response, may itself lead to further oxidative damage in the cell and activation of proteases, which cleave MqsA leading to persister formation (Wang *et al.*, 2011). Another toxin Hha is under positive control of MqsR. Moreover, Hha induces other toxins including RelE, YoeB, YafQ and also activates specific proteases like ClpP and ClpX. This suggests that Hha activates toxins and promotes degradation of antitoxins, which increases persister cell formation (Kim & Wood, 2010). The expression of Hha influences biofilm development by decreasing motility and promoting cell

aggregation during biofilm formation (Barrios *et al.*, 2006). Hence, the Toxin-Antitoxin systems acts as an important regulators of the switch from a planktonic (highly motile) to a biofilm lifestyle (non-motile) (Wang *et al.*, 2011). Expression of the toxin YafQ cleaves the *rpoS* transcript and therefore reduce RpoS signalling (Prysak *et al.*, 2009). Overexpression of YafQ increases resistance to antibiotics such as ampicillin and ciprofloxacin. The increase in resistance is due to cleavage and reduction in tryptophanase (*tnaA*) transcript leading to reduction in indole signaling, a negative regulator of persister cell formation (Hu *et al.*, 2015).

1.2 Unbiased approaches to study gene expression

1.2.1 Introduction

Functional genomics includes study of the expression levels of RNA and proteins in a cell by focusing on dynamic processes such as gene transcription, translation, and regulation of gene expression as well as protein-protein and protein-DNA interactions. A genome-wide approach is usually taken to study the gene regulation than the more traditional "gene-by-gene" approach (Table 2).

Table 2: Genome wide approaches for transcriptomic analysis.

Techniques	Uses	Advantages	Disadvantages	References
Transgenesis of Reporter gene	"Gold standard" and accurate method for functional analysis of regulatory elements	• Low cost • Gene expression is easily detectable	• Regulatory elements are widely dispersed through the genome that may cause some difficulties in detection • Genetic engineering is employed	(Loots, 2008, Andersen *et al.*, 1998, Uliczka *et al.*, 2011, Hsiao & Zhu, 2009)
Serial Analysis of Gene Expression (SAGE)	Direct and quantitative method for snapshot of mRNA population in a sample of interest	• *Prior* knowledge about the gene sequences is not required. • SAGE library requires a small amount of RNA as input. • Simple data analysis.	• Low-throughput • Difficulty to construct tag libraries • Cost and time to perform so many PCR and sequencing reactions • Limited by total number of tags sequenced	(Hu & Polyak, 2006)
Microarray	Well-studied high throughput and quantitative method for gene expression studies	• High throughput • Based on fluorescence (no need of radioactive probes)	• Knowledge of sequences required • Relies on annotated genome • Limited by hybridization signal	(Kostić *et al.*, 2007, Zhou, 2003)

Techniques	Uses	Advantages	Disadvantages	References
		• Fast and easy to perform	• Genes with low expression may not be detected high signal-to-noise ratio	
RNA sequencing	Direct, quantitative and high throughput method. Appropriate for gene, transcripts (including alternative gene spliced transcripts) or allele-specific expression identification	• Prior knowledge about the genomic features is not required • Ability to detect novel transcripts • Wider dynamic range • Higher specificity and sensitivity • Simple detection of low abundance transcripts	• Sequencing depth may affect dynamic range and reproducibility. • Complicated and multiple ways of data analysis • Expensive technique	(Croucher & Thomson, 2010)
ChIP-seq or ChIP-chip	Method to study protein-DNA interaction. Fast and well-studied	• Compatible with array-or sequencing-based analysis • Possible to perform genome-wide analysis	• Generates large dataset • Expensive technique	(Shendure & Aiden, 2012, Wong *et al.*, 2017)
Transposon mutagenesis high throughput sequencing (Tn-seq)	High throughput parallel sequencing for fitness and genetic interaction studies in microorganisms	• Capacity to screen an entire library in a single infectivity experiment • Does not require isolation and characterization of individual Tn mutant clones • Easier application to in vitro screening methods	• Lack of plasmid content information; low infectivity may be related to plasmid loss • Tn mutants of interest would have to be re-isolated for further study • Possible bottleneck effects (non-uniform recovery of organisms) may necessitate use of	(Van Opijnen *et al.*, 2009, Lin *et al.*, 2014, Goodall *et al.*, 2018)

Techniques	Uses	Advantages	Disadvantages	References
			large numbers of animals or cultures • Relatively high minimum analysis cost	

Next Generation Sequencing (NGS) is used in genomics to sequence the complete genome of organisms. This helps to elucidate DNA mutations, ranging from single nucleotide polymorphisms to large gene deletions or insertions (Heather & Chain, 2016). At the transcriptomic level, DNA microarrays (Schulze & Downward, 2001) and RNA-sequencing (Wang *et al.*, 2009) are the most established and recent techniques used for profiling the gene expression of organisms and are discussed below in detail. Mass spectrometry (MS) combined with 2D gel electrophoresis (2-dimensional gel) can be used to study the functions of proteins as well as to quantify the protein abundance. Multidimensional Protein Identification Technology (MudPIT) is a widely used technique that digests proteins into peptides and then separates them by two-dimensional chromatography based on charge and hydrophobicity, and are subsequently analyzed by MS (Graves & Haystead, 2002). Protein-protein interactions can be examined by protein microarray and Affinity purification technique followed by MS. Protein-DNA interactions are widely determined by combining chromatin immunoprecipitation assay with sequencing (ChIP-seq) or microarray (ChIP-chip) technology. These techniques help to identify the genome-wide DNA binding sites for transcriptional factors and other DNA binding proteins (Pepke *et al.*, 2009).

1.2.2 Microarray

After the development of DNA microarray techniques, gene expression analysis was revolutionized by its capacity to examine thousands of RNA products at once. There are three

basic types of microarrays these are spotted arrays, in-situ synthesized arrays, and self-assembled arrays (Bumgarner, 2013). Oligonucleotide arrays is a variant of in-situ synthesized arrays. This technique has many advantages over other types of microarray including high specificity, good reproducibility and precise measurements of gene expression. An oligonucleotide array can be a single-channel array that is hybridized with only one sample, and therefore generates absolute expression levels (Jaluria *et al.*, 2007). A microarray chip is little more than a glass microscope slide with thousands of spots on it (often referred as chip). Each spot is a unique nucleotide sequence that serves as a probe for a specific gene. The different DNA fragments (referred to as oligonucleotides) are arranged in orderly rows and columns such that the identity of each fragment is known through its location on array. The size of oligonucleotides can be short (15-25 nucleotides) or long (50-100 nucleotides). The use of longer oligonucleotides may increase the specificity of hybridization and sensitivity of detection (Schulze & Downward, 2001). The principle behind microarrays is measuring signal intensities of hybridization. The RNA sample, extracted from a desired sample, is transcribed into its complementary DNA (cDNA) and labeled with either a fluorescent dye (fluorophore) or a radioactive isotope. Single-channel array can also be used which include single fluorophore labelling, for instance, during a time course analysis to study different growth phases of bacteria. Two different fluorophores can also be used for simultaneous detection and comparison of known samples termed as Dual-channel microarray, for instance, when comparing bacterial growth in rich media versus minimal media. The most commonly used fluorescent dyes for cDNA labeling include Cy3 (green fluorescent) and Cy5 (red fluorescent). The labeled targets are allowed to hybridize to probes, which will undergo competitive binding between the different samples to the corresponding array probe. The chip is washed and scanned using a laser confocal microscope to excite hybridization fluorophores. The

relative fluorescence of each spot is detected and recorded (Figure 1). The generated data can be further analyzed to determine the gene expression in a particular sample (Jaluria *et al.*, 2007). Microarrays are widely used to study the mRNA gene expression. Moreover, the use of microarrays in laboratories has expanded to study toxicology, evolutionary biology, drug development, cellular physiology and stress responses, and forensic science (Miller & Tang, 2009).

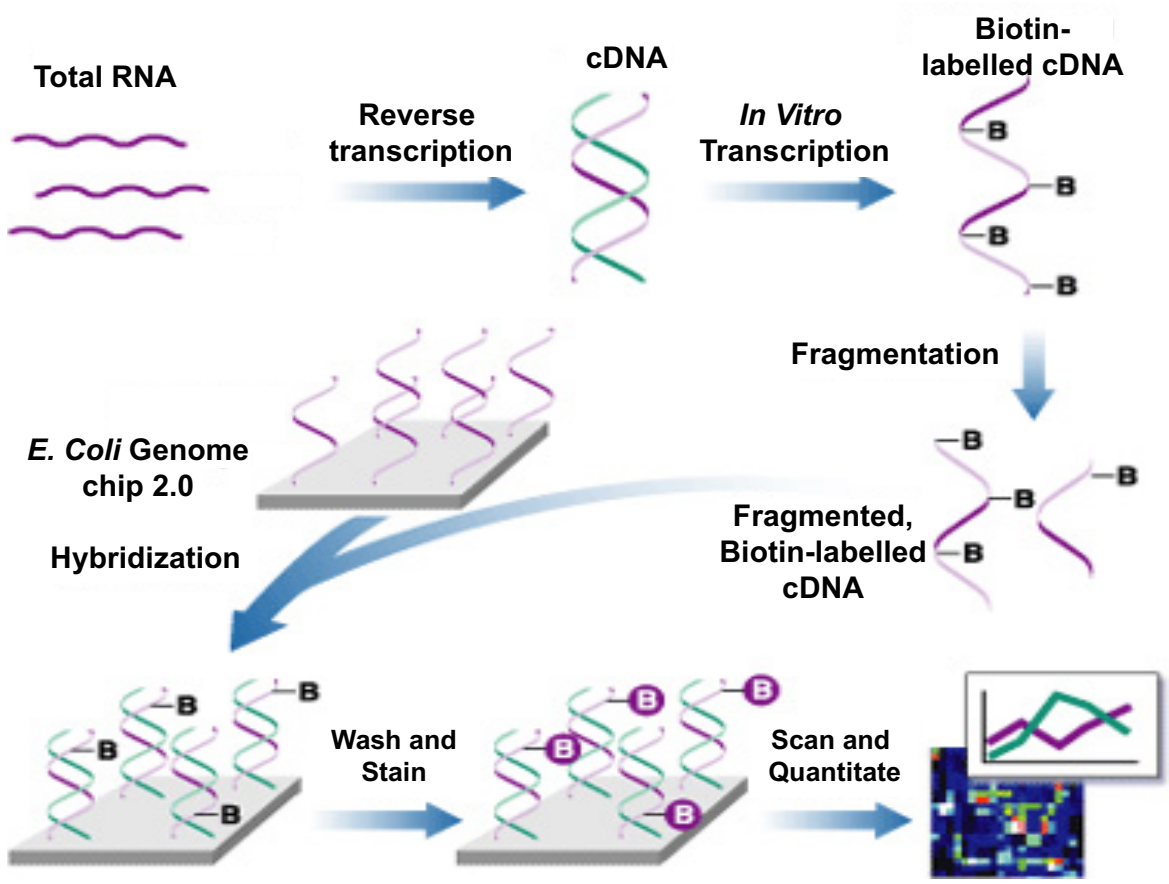


Figure 1: Overall method to perform microarray technique.
The total RNA is extracted from samples and reverse transcriptase is used to copy the RNA into stable ds cDNA. In microarray, the ds cDNA is fragmented labelled with biotin or florescent dye. The labelled fragments bind to an ordered array of complimentary oligonucleotides, and measurement of fluorescent intensity across the array indicates the abundance of a predetermined set of sequences. Image from (Ryan et al., 2004)

Affymetrix GeneChip arrays are high-density oligonucleotide expression arrays and are used in the current study. The mRNA sequence of a gene is represented by a probe set composed of 11-20 probe pairs. Each probe pair is composed of a perfect match (PM) probe, a 25-base pair DNA copy of a section of the mRNA sequence of interest, and a mismatch (MM) probe, that is created by changing the 13th base pair of the PM probe with the intention of measuring non-specific binding (NSB). The RNA samples are labeled and hybridized with arrays, subsequently scanned by an Affymetrix scanner which generates raw optical/pixel intensities of each spot and stores that data in DAT format files. The DAT format files are further processed through Gene Chip Operating Software (GCOS) by converting the DAT file into CEL file format which stores the results of the intensity calculations on the pixel values of DAT file. The CEL files can be further used for downstream analysis which transforms intensity levels into expression values. These intensities represent the amount of hybridization for each oligonucleotide probe. However, a part of the hybridization is non-specific, and the intensities are affected by optical noise. Therefore, the observed intensities need to be adjusted to provide accurate measurements of specific hybridization. The final step is to combine the 11-20 probe pair intensities, after background adjustment and normalization, for a given gene to define a measure of expression.

Several algorithms can be used to analyze microarraye data such as LOESS, dChip (Li & Wong, 2001), MAS5 (Hubbell *et al.*, 2002), PLIER (Xing *et al.*, 2006), RMA (Irizarry *et al.*, 2003), and GCRMA (Naef & Magnasco, 2003) for preprocessing and normalization. There are many theoretical and empirical advantages and disadvantages to the different steps in each processing pipeline. Although a single method is not superior to others, it is concluded that the efficiency of the method is affected by the nature of the study (Verhaak *et al.*, 2006). Preprocessing of microarray raw data is a three-step process for Affymetrix data that results in the summed

normalized signal intensity measurements for each gene. The first crucial step is called background correction, which filters the background noise from the data generated due to non-specific and false binding. Secondly, normalization is applied to enhance the comparison between different data from different microarray experiments by adjusting and scaling the main characteristics of the data, including mean/median, distribution and/or standard deviation. The last step is a summarization of the normalized values. After the normalization of the signal intensities of each probe, the values are collected and summed into a single signal intensity value. Although some methods might have a different order or extra steps during preprocessing (Schuster *et al.*, 2007). Each of the above mention algorithms employs different methods for background correction/subtraction, signal normalization and probe set summarization.

GCRMA (GeneChip Robust Multiarray Averaging) algorithm is used in this study which is an enhanced version of RMA (Robust Multiarray Average) algorithm that uses GC content information of each nucleotide to calculate binding efficiency and thus, signal intensity. Since the strength of G-C hybridization is stronger than A-T, the GC content of an oligonucleotide affects the binding tendency of each oligonucleotide pair after washing the arrays (Naef & Magnasco, 2003). So for background correction, GCRMA background correction method is applied (Lim *et al.*, 2007). Normalization and summarization steps are same as the RMA method. Quantile normalization is used which is a linear method for array-wise adjustment that scales the data across the arrays in quantiles (Bolstad *et al.*, 2003). Lastly, median polish, a summarization method, is used for getting a single signal intensity value for a transcript from multiple oligonucleotides (Lim *et al.*, 2007). Median polish minimizes the residual log error. As a result, different signal intensities are transformed into one average distribution.

1.2.3 RNA-sequencing

The advent of high-throughput sequencing-based methods has changed the way transcriptomes are studied. The RNA-sequencing technique includes direct sequencing of complementary DNA using next-generation sequencing technologies. Next-generation sequencing technologies have many advantages including single base pair resolution, a low background signal, a large dynamic range of expression over which transcripts can be detected, higher levels of reproducibility, and small sample quantity (Wang *et al.*, 2009). During RNA-sequencing (Figure 2) rRNA is depleted from total RNA extracts of a sample. The mRNA enriched sample is then converted into double-stranded cDNA using reverse transcriptase with random or oligo (dT) primers and the full-length cDNA is fragmented and ligated to an adaptor for amplification by PCR (Nagalakshmi *et al.*, 2010). Each amplified molecule with amplification is then sequenced in a high-throughput manner to obtain short sequences from one end (single-end sequencing) or both ends (paired-end sequencing). The reads are typically 30-400bp, depending on the DNA-sequencing technology used (Wang *et al.*, 2009). High-throughput technologies such as Illumina, ABI and Roche 454 are used for bacterial RNA-seq experiments (Croucher & Thomson, 2010). Following sequencing, the actual RNA-seq data analysis has many variations, as the applications of the technology is diverse. One of the challenges associated with RNA-seq analysis is that, though there are numerous tools available (Han *et al.*, 2015) that support several aspects required for analysis, most of these tools are designed primarily for use with eukaryotic genomes. Bacterial RNA-seq faces different challenges from eukaryotic RNA-seq including overlapping of genes; therefore, distinguishing the start of one gene transcript from the end of another adds a layer complexity. The prevalence of polycistronic messages further complicates bacterial transcript assembly. Moreover, the models for eukaryotic RNA gene analysis are not suitable for small

regulatory RNA (sRNA). The major steps for typical RNA-seq analysis include quality control check, performing read alignments that map the resulting reads either to a reference genome or assembling the overlapping reads without the reference genomic sequence (*de novo* assembly), obtaining raw counts and detecting differential gene expression. Many softwares can be used to produce a genome-scale transcription map that consists of both the transcriptional structure and level of expression for each gene (Conesa *et al.*, 2016).

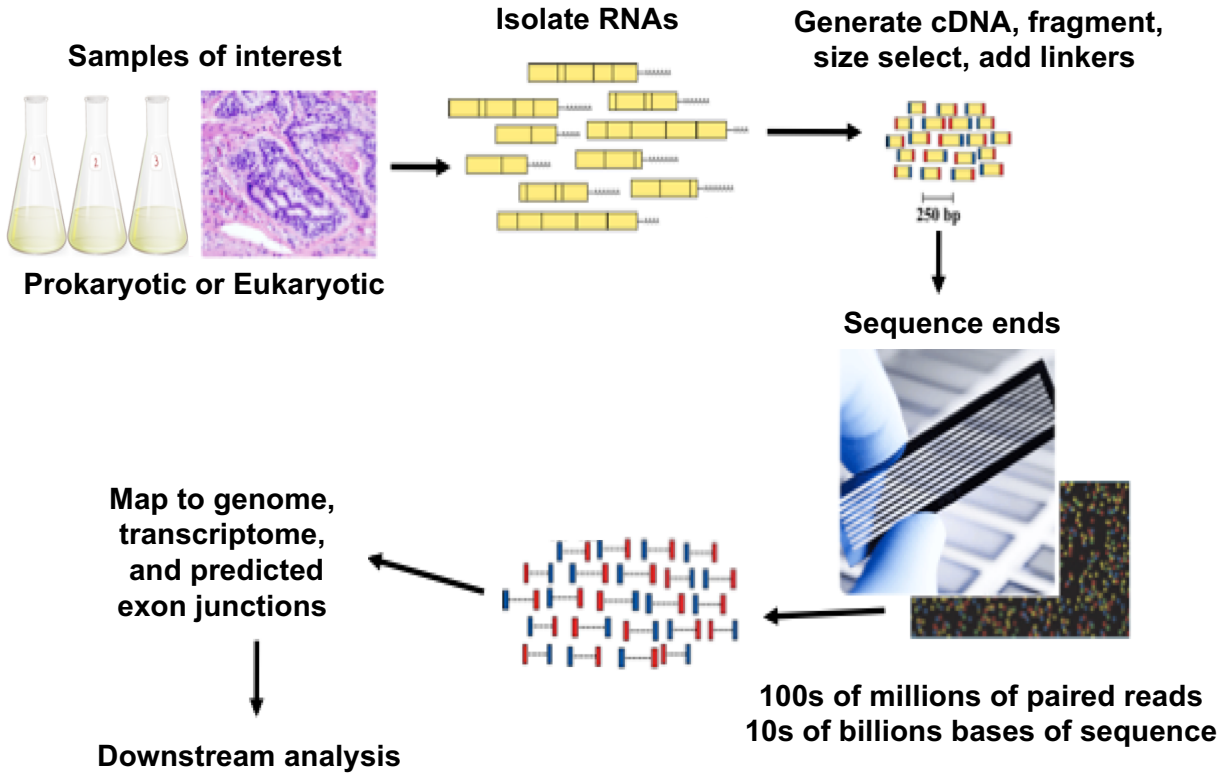


Figure 2: Overall procedure for performing RNA-sequencing.
A typical RNA-sequencing workflow includes the isolation of RNA from sample of interest, generation of sequencing libraries, use of a high-throughput sequencer to produce hundreds of millions of short paired-end or single-end reads, alignment of reads against a reference genome or transcriptome, and downstream analysis for expression estimation, differential expression, isoform discovery, and other applications. Image from (Griffith *et al.*, 2015)

Quality control for the raw reads includes the analysis of sequence quality, GC content, the presence of adaptors, overrepresented k-mers and duplicated reads to detect sequencing errors, PCR artifacts or contaminations. Acceptable duplication, k-mers or GC content levels are experiment- and organism-specific, but these values should be homogeneous for samples in a particular experiment. FastQC (Andrews, 2014) can be used to perform these analyses. Trimmomatic (Bolger *et al.*, 2014) can be used to discard low-quality reads, trim adaptor sequences, and eliminate poor-quality bases. Processed raw reads are typically mapped to a genome. An important mapping quality parameter is the percentage of mapped reads, which is a global indicator of the overall sequencing accuracy. *De novo* mapping leads to the discovery of new and unannotated transcripts. The most common application of RNA-seq is to estimate differential gene expression. This is primarily based on the number of reads that map to each transcript sequence. The simplest approach to quantification is to aggregate raw counts of mapped reads using programs such as HTSeq count (Anders *et al.*, 2015) or FeatureCounts (Liao *et al.*, 2013). This gene-level quantification approach utilizes a gene transfer format (GTF) file (Bioinformatics, 2016) containing the genome coordinates of exons and genes, and often discard multireads. Raw read counts alone are not sufficient to compare expression levels among samples, as these values are affected by factors such as transcript length, total number of reads, and sequencing biases. For instance, longer genes have a greater likelihood of being detected as of more read counts, while in parametric method for differential expression analysis tend to yield more DE expressed genes as sequencing depth is increased. There are different normalization methods to overcome many affected factors, includes Total counts(TC), Median (Med) (Dillies *et al.*, 2012), Upper quantile (UQ) (Bullard *et al.*, 2010), Trimmed Mean of M-values (TMM) (Robinson & Oshlack, 2010), DESeq (Anders & Huber, 2010), Quantile (Q) (Bolstad *et al.*, 2003),

FPKM (Fragments Per Kilobase Million), TPM (Transcripts Per Kilobase Million) and RPKM (Reads Per Kilobase of transcripts per Million mapped reads) (Pachter, 2011). The measure RPKM is a within-sample normalization method that will remove the feature-length and library-size effects and used to represent gene expression value. Correcting for gene length is necessary for correctly ranking gene expression levels within the sample to account for the fact that longer genes accumulate more reads but is not necessary when comparing gene expression changes within the same gene across all samples. So RPKM, FPKM, and TPM normalization methods cannot be used for the differential expression analysis, as it does not consider the most important factor, sequencing depth for comparison among the samples (Bullard *et al.*, 2010). Differential expression analysis requires the gene expression values should be compared among samples. The normalization methods that take this into account are TMM, DESeq, and Upper Quartile which ignore highly variable features. DESeq2 estimate the variance in RNA-seq data and test for differential expression. It performs an internal normalization where geometric mean is calculated for each gene across all samples. The counts for a gene in each sample is then divided by this mean. The median of these ratios in a sample is the size factor for that sample. This procedure corrects for library size and RNA composition bias and make data comparable across all samples (Love *et al.*, 2014). In addition to all, it is also crucial to assess the global quality of the RNA-seq dataset by checking on the reproducibility among replicates and for possible batch effects. Reproducibility among technical replicates should be generally high (Mortazavi *et al.*, 2008), but no clear standard exists for biological replicates, as this depends on the heterogeneity of the experimental system. If gene expression differences exist among experimental conditions, it should be expected that biological replicates of the same condition will cluster together in a Principal Component Analysis (PCA).

1.3 Project rationale

In a previous study to identify genes controlled by RpoS, we found that a large fraction of *E. coli* genome continues to increase in expression during prolonged starvation that does not require RpoS (Schellhorn HE, 1998) (Figure 3). This suggests that in addition to RpoS, other growth phase-dependent regulatory mechanisms may control prolonged stationary phase gene expression. These genes may be important for adaptation and survival under nutrient-limiting conditions. Studying gene expression for prolonged-incubation phase cultures is difficult as RNA yield for old culture is low compared to exponentially growing cultures, which may be due to ribosomal degradation in long-term cultures. Hence, for quality control purposes we used both rRNA depleted (RNA-seq) and non-rRNA depleted technique (microarray) to examine the transcript abundance during prolonged-incubation phase.

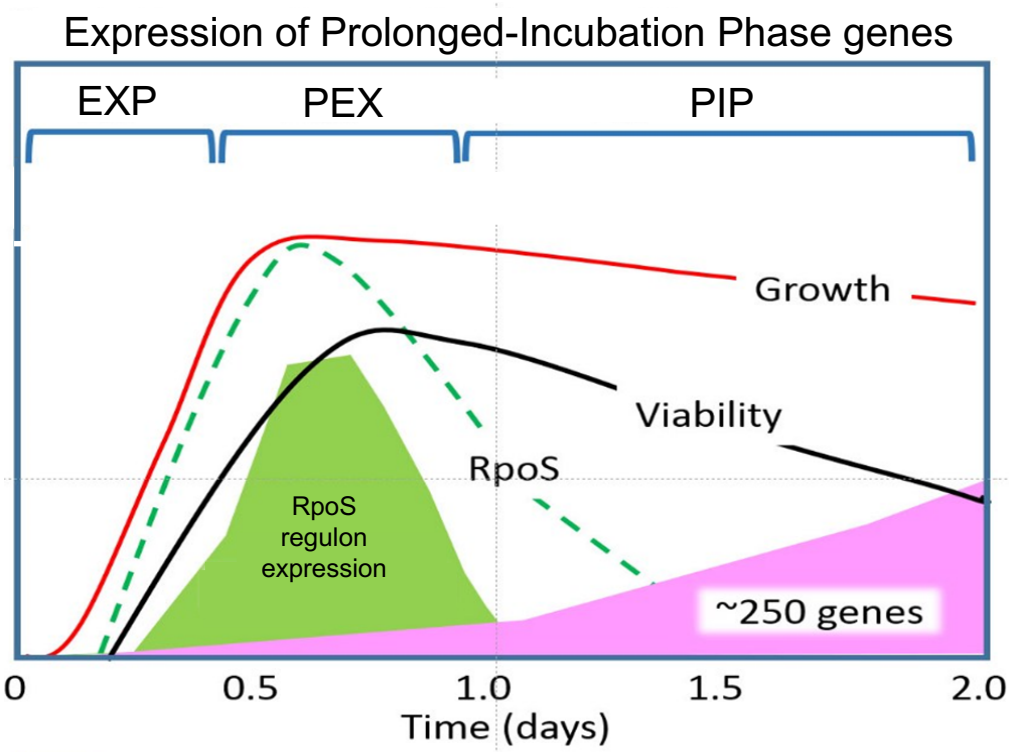


Figure 3: Predicted gene expression of *E. coli* for distinct growth phases.
Exponential phase (EXP), Post exponential phase (PEX) and Prolonged-incubation phase (PIP) (Figure from Schellhorn lab)

1.4 Goals

1) To identify genes that are overexpressed during prolonged-incubation phase of *E. coli* using transcriptome-based technologies.

2) To determine whether the overexpressed genes are RpoS-dependent or independent during prolonged-incubation phase.

3) To determine if the genes that are overexpressed during prolonged-incubation phase are distinct from the stationary phase regulon member.

2 METHODS AND MATERIALS

2.1 Growth conditions

Overnight cultures of the MG1655 strain (WT) and isogenic *rpoS* mutant (Δ*rpoS*) strain of *E. coli* were grown from single, independent colonies in LB broth (LB-Miller, 10 g of peptone, 5 g of yeast extract, and 10 g of NaCl). Overnight cultures were sub-cultured to 1:10,000 dilution in fresh LB and grown at 37 °C with shaking at 200 rpm in 50 ml of LB in 250 ml flask with aeration.

2.2 RNA extraction and quality check

Total RNA was isolated from the sub-cultured cells of both the strains at exponential phase (OD600 = 0.3), early stationary phase (OD600 = 1.5), and prolonged-incubation phase (24h and 48h) using the Norgen Total RNA Purification Kit. Quality of the isolated RNA was analyzed using Nanodrop (Thermo Scientific), the OD260/280 ratio was taken and using Invitrogen Qubit (Q32855). RNA was DNase treated and re-purified using RNA Clean and Concentrator kit (Zymo Research). Three biological replicates were collected for each phase. Calculated RNA yield and integrity were checked on 0.8% agarose gel stained with ethidium bromide (Appendix 1). The extracted RNA sent for RNA-sequencing (rRNA-depleted) and microarray (non-rRNA depleted) analysis.

2.3 Library preparation and analysis of RNA-seq data

The RNA samples were rRNA depleted using Ribo-Zero kit (Illumina). RNA-Sequence libraries were prepared using TruSeq RNA Sample Prep kit (Illumina) and sequenced on the Illumina HiSeq 2000 platform with single-end reads and read lengths of 50 nt at the McGill University and Génome Québec Innovation Centre (Montreal, Canada). On average ~2 million reads were obtained per cDNA library after sequencing. The preliminary quality control for files containing single ended 50 bp illumina reads were checked using FastQC (Andrews, 2014). The

analysis indicated that Truseq adaptors had already been trimmed and low-quality reads were removed by the sequencing facility. Reads were mapped to the NCBI K12 reference genome (NC_000913.2 *Escherichia coli* str. K-12 substr. MG1655) using Bowtie2 (Langmead & Salzberg, 2012). The percentage of mapped reads was assessed (Appendix 2). The reads mapped to each gene was counted with HTseq 0.9.1 (Anders *et al.*, 2015). Reads that were uniquely aligned to each gene were tabulated from each replicate separately. Differentially-expressed genes were estimated using DESeq2 package available under the open-source Bioconductor suite of programs (Reimers & Carey, 2006). DESeq2 estimate the variance in RNA-seq data to test for differential expression (Anders & Huber, 2010). As an input, DESeq2 accepts a table of raw read counts for each gene from different biological replicates, and estimates the differentially-expressed genes using a negative binomial distribution (Love *et al.*, 2014). DESeq2 performs an internal normalization where geometric mean is calculated for each gene across all samples. The counts for a gene in each sample is then divided by this mean. The median of these ratios in a sample is the size factor for that sample. This procedure corrects for library size and RNA composition bias. The program-generated *p*-values were used to determine the significance of the differential levels of gene expression based on the Benjamini-Hochberg correlation, with a false-discovery rate of < 5%. In both strains, the transcripts determined to be significantly altered if the fold change ≥ 4.0 with FDR adjusted $p \leq 0.05$. Principal component analysis was performed to check the clustering among replicates for all time points. Functional enrichment of differentially-expressed genes was determined using online EcoCyc pathway tools (Karp *et al.*, 2014). Ecocyc is a database available at https://ecocyc.org/ that describes the genome and the biochemical machinery of *E. coli* K-12 MG1655. It also facilitates the analysis of high-throughput data including gene-expression and metabolomics data through different tools. The data can also be visualized on a metabolic map

diagram, complete genome diagram, or regulatory network diagram. Functional enrichment analysis was assessed using Fisher's-exact statistical test together with Bonferroni Correction method (Ashburner *et al.*, 2000). If a GO term in a test gene set showed a corrected p-value ≤ 0.05, the GO term (function) was considered to be significantly overrepresented.

2.4 Microarray analysis

Single-channel microarray analysis was performed at The Centre for Applied Genomics (TCAG), The Hospital for Sick Children, Toronto. RNA samples were each labelled with Biotin Allonamide Triphosphate (single color), individual samples hybridized to Affymetrix *E. coli* Genome 2.0 GeneChip and the arrays were scanned at Affymetrix GeneChip Scanner 3000. The generated microarray CEL files were preprocessed in R software using Bioconductor's *affy* package and were normalized using the GCRMA (GeneChip Robust Multiarray Averaging) method. GCRMA is an improved form of RMA (Robust Multiarray Average) that include optical noise and Non-Specific Binding (NSB) to adjusts for background intensities in Affymetrix array data. GCRMA converts background adjusted probe intensities to expression values using the same normalization and summarization methods as RMA. The resulted log transformed, and normalized datasets were used for differential expression analysis. Differential expression analysis was performed also in R software, using the *limma* package (lmFit and eBayes methods) (Smyth & Speed, 2003). The eBayes (Smyth, 2004) test was used to generate P values and to determine the significance of the differential levels of gene expression based on the Benjamini-Hochberg correlation, with a false-discovery rate of < 5%. In both strains, the transcripts determined to be significantly altered if the fold change ≥ 4.0 with FDR adjusted $p \leq 0.05$.

2.5 RT-qPCR for validation of gene expression data

Expression of representative identified genes from the different functional groups was quantified by quantitative real-time PCR (qPCR) using Bio-Rad CFX96 Real-Time PCR System. Optimized primers were designed using Blast-NCBI primers design (length 18-22 nucleotides and predicted annealing temperature ranging from 55 to 60 °C) (Appendix 5) to amplify about 70-150 base pairs of the target genes. Reverse transcription was performed on 500 ng of each RNA samples for all growth phases with random hexamer primers using iScript cDNA Synthesis Kit (Biorad) according to the manufacture's protocol. mRNA levels were quantified using SsoFast Evagreen Supermix and CFX-96 Real Time PCR system. The *rrsA* gene, encoding the 16S ribosomal RNA was used as a reference gene for normalization. The same RNA sample were included in the PCR reaction as a negative control to test for genomic contamination. The RT-qPCR assays were conducted in triplicate biological RNA samples. Specificity and efficiency of amplification of each primer pair was verified by constructing a standard curve of amplification on a serial dilution of the purified *E. coli* genomic DNA template to confirm that each of the assays were conducted in the linear range and the slope of the threshold cycle Ct when plotted against the dilution were within the acceptable range for all the assays (Appendix 3).

3 RESULTS AND DISCUSSION

3.1 Comparison of RNA-sequencing and Microarray data

3.1.1 Effect of rRNA removal procedure

Bacterial transcriptomes contain protein-coding RNA, transfer (t)RNA, transfer messenger (tm)RNA, small regulatory (s)RNA, and ribosomal (r)RNA. Ribosomal RNA accounts for more than 85% of prokaryotic cellular RNA content (Karpinets *et al.*, 2006), which can impede the analysis of mRNA transcripts, with ≥80% of library cDNAs mapping to rRNA in the absence of selection procedures (Van Vliet, 2010). RNA-sequencing technique removes rRNA and constructs cDNA libraries from rRNA depleted samples. Several methods developed for rRNA depletion including exonuclease treatment, polyadenylation (Shi *et al.*, 2009, Wendisch *et al.*, 2001), electrophoretic size separation (McGrath *et al.*, 2008), and subtractive hybridization capture the rRNA (Su & Sordillo, 1998). In contrast, microarray analysis utilizes the hybridization of total sample including RNA content for gene expression determination. To test the effect of ribosomal depletion on gene expression analysis of *E. coli* cultures, we performed RNA-seq and microarray analyses of exponential, early stationary and prolonged-incubation phase cultures.

The subtractive hybridization process, which is included in the workflow of several commercial kits, is the most common choice for rRNA depletion for prokaryotic RNA-seq analyses. The Ribo-Zero kit we used employs biotinylated rRNA capture probes for 16S, 23S and 5S rRNA encoded by seven operons (*rrnA, rrnB, rrnC, rrnD, rrnE, rrnG* and *rrnH*). These probes specifically hybridize to rRNA molecules and are subsequently captured by magnetic beads which are removed from the sample. Recently, a comparative evaluation of rRNA depletion procedures of bacterial biofilm and mixed pathogen culture transcriptomes suggested that the Ribo-Zero kit exhibited the highest efficiency compared to other commercially available kits such as the Ambion

MICROBExpress™ Bacterial mRNA Enrichment Kit and the Life Technologies RiboMinus Transcriptome Isolation Kit (Petrova *et al.*, 2017).

The current study includes RNA samples from old cultures which includes one and two-day old planktonic cultures. The old cultures had low RNA yield compared to exponentially growing cultures. The reason for this low RNA yield is still unknown. Another study found that the lower proportions of mRNA reads are consistently sequenced from biofilm cultures that might have higher rRNA:mRMA ratios than those of planktonic cultures (Dötsch *et al.*, 2012). This suggests that the efficiency of rRNA removal strongly depends on the culture conditions. To measure the rRNA removal efficiency, we assessed the presence of 16S, 23S and 5S rRNA transcript abundance in RNA-seq and microarray data. As expected, RNA-seq data showed low abundances of 16S, 23S and 5S rRNA transcripts compared to the signal intensities of 16S, 23S and 5S rRNA probe (Table 3). The analysis verifies that rRNA depleted samples have low counts for rRNA genes compared to non-depleted samples. The ribosomal associated protein coding genes had high transcript abundance (For instance, thousands of reads were mapped) compared to 16S, 23S and 5S rRNA (where the number of mapped reads were in the hundreds). This justifies the 80 – 90% of rRNA removal efficiency of the Ribo zero kit.

Furthermore, the ribosomal associated protein coding genes had low transcript abundance during prolonged-incubation (specially at 48h of incubation) in both RNA-seq and microarray data (Figure 4). This suggests that ribosomal protein-coding genes are decreased in expression during prolonged-incubation. This may be the possible reason for low RNA yield in old cultures. A rapid accumulation of ppGpp may impede rRNA synthesis and subsequently growth arrest (Cashel, 1996). However, the precise reason is still illusive.

Table 3: Transcripts abundance of *rrn* operon in RNA-sequencing and microarray data.

	WT				Δ*rpoS*			
Genes	0.3 OD	1.5 OD	24h	48h	0.3 OD	1.5 OD	24h	48h
RNA-sequencing data								
16S rRNA								
rrsA	9.3	8.7	9.8	8.1	9.4	8.3	8.4	9.9
rrsB	2.5	3.2	3.0	2.5	4.2	3.2	4.6	4.2
rrsC	6.5	6.4	6.6	6.2	6.7	5.6	6.2	7.5
rrsD	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
rrsE	4.8	4.2	4.2	4.7	4.7	4.5	5.8	5.9
rrsG	4.5	4.5	5.0	4.9	3.9	3.7	4.5	5.6
rrsH	5.7	6.2	5.5	7.9	8.7	10.3	10.5	9.0
23S rRNA								
rrlA	10.4	9.8	10.7	8.0	10.3	8.9	7.6	10.1
rrlB	2.4	2.9	4.2	6.2	2.0	1.6	4.6	3.3
rrlC	8.2	9.6	10.6	8.4	7.1	8.6	8.4	10.0
rrlD	3.5	3.8	4.2	4.5	2.0	2.2	4.6	3.9
rrlE	8.3	8.0	8.2	5.4	8.0	7.7	5.7	8.2
rrlG	9.6	8.3	8.3	5.9	9.2	7.5	5.4	7.9
rrlH	2.3	4.9	5.9	5.3	2.3	3.2	4.4	4.4
5S rRNA								
rrfA	5.9	5.7	4.3	5.0	5.1	6.5	6.8	6.6
rrfB	7.5	8.4	8.2	6.7	7.4	7.4	7.4	7.8
rrfC	6.8	6.5	6.4	6.3	6.7	6.1	5.4	5.7
rrfD	6.9	6.4	6.2	5.3	6.5	7.3	6.6	7.4
rrfE	0.0	0.0	0.0	0.0	1.7	2.4	3.8	2.8
rrfF	5.2	5.5	6.1	6.1	4.8	5.7	5.6	5.1
rrfG	5.4	6.6	8.7	7.3	5.2	5.6	6.0	6.1
rrfH	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Microarray data								
16S rRNA probe	13.5	13.9	15.0	15.7	13.3	13.9	15.0	15.6
23S rRNA probe	13.0	13.1	14.2	14.6	12.8	13.4	14.8	14.5
5S rRNA probe	12.0	12.2	12.9	13.7	11.5	11.4	13.4	11.5

(The maximum log2RPKM value in each growth phase is approximately 15.5, which is high compared to rRNA transcript values in RNA-seq data. The maximum log2 signal intensities in each growth phase is approximately 15.2, which is close to rRNA probe intensity values in microarray data)

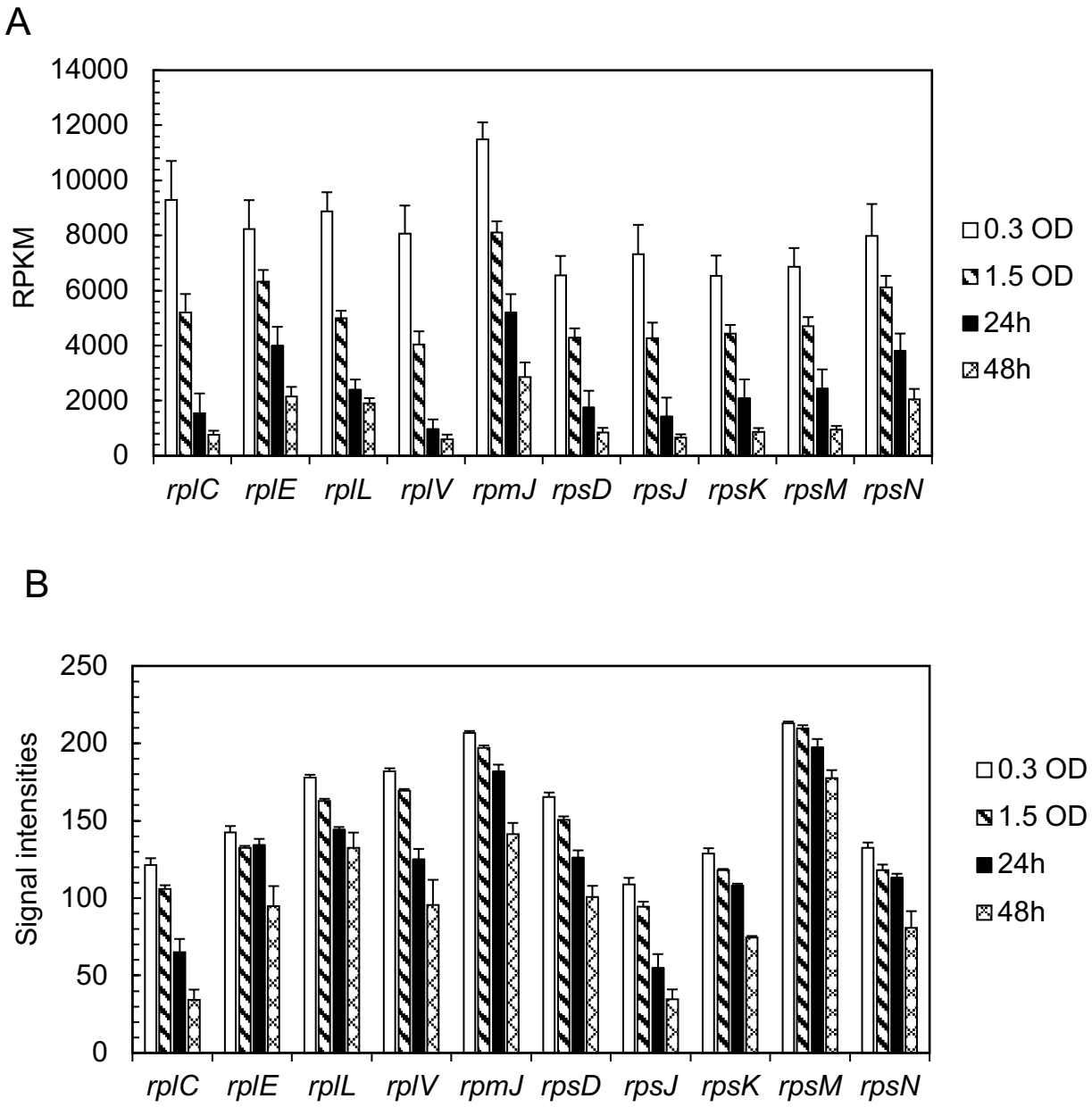


Figure 4: Ribosomal associated protein-coding transcripts were in low abundance during prolonged-incubation.
There are several ribosomal associated protein-coding genes, however, only few top most down-regulated are shown from RNA-sequencing (A) and microarray data (B). Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

3.1.2 Correlation analysis

Considering the observed differences in microarray vs RNA-seq, we proceeded to compare the expression data using Pearson correlation analysis. For RNA-seq data, RPKM (Reads Per Kilobase of transcript per Million mapped reads) value was used. RPKM is generated by dividing the total mapped reads by gene length and the total number of mapped reads of a sample. It is a within-sample normalization method that removes the feature-length and library-size bias. For microarray data, the GCRMA method was used, which converts the probe intensities to log2 transformed signal intensities. To make the transcriptome profiles comparable between the two platforms, the generated RPKM values were log2 transformed. The transcripts annotated in both techniques were used for Pearson correlation analysis (Appendix 3). The processed data from both the platforms was correlated and compared in R software using the *corrplot* package. The resulting correlation was mapped as a scatter plot, with the average numbers of counts from RNA-sequencing against the normalized fluorescence intensities, from microarray, for each gene in the WT (Figure 5) as well as in the Δ*rpoS* (Figure 6). Pearson correlation coefficient between the platforms ranges from 0.6 – 0.8 ($p \leq 0.01$) and the level of significance was checked by *t*-test. This is in agreement with the previous reports that expression levels measured by microarray and RNA-seq have correlations ranging between 0.6 and 0.8 for prokaryotic datasets (Zhao *et al.*, 2014, Nookaew *et al.*, 2012). This analysis showed that the gene expression data is highly correlated between the techniques. Despite, high correlation between the techniques few low abundances transcripts were discrete in both techniques, which is consistent with the other study (Chen *et al.*, 2017). Moreover, the use of different normalization methods in both platforms may also affect the correlation analysis.

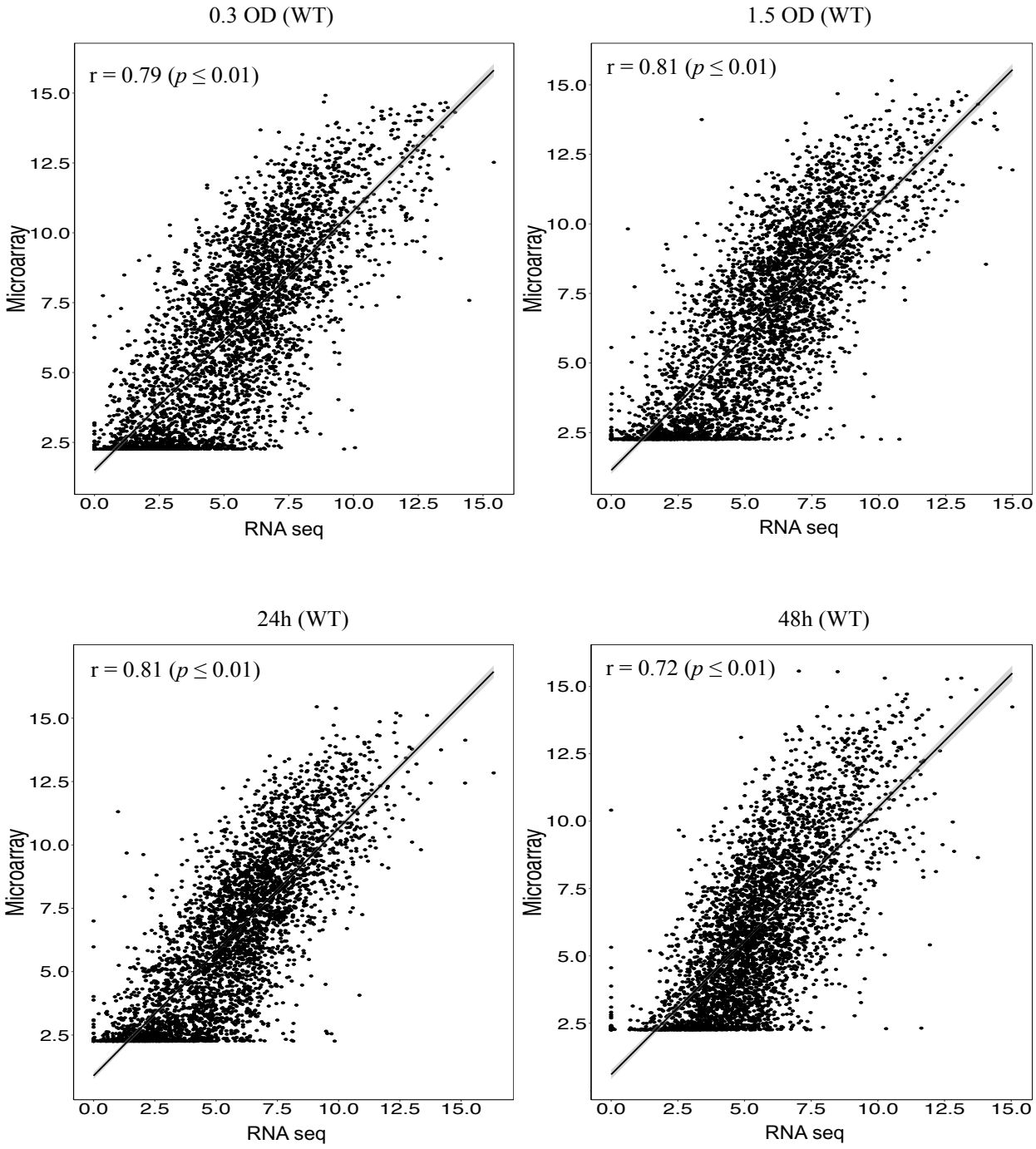


Figure 5: Pearson correlation analysis of RNA-seq and microarray data for WT.
The relationship between the expression profiles generated by both platforms is depicted as a linear regression line. Pearson correlation coefficient represented by r value, and *p*-value shows the level of significance.

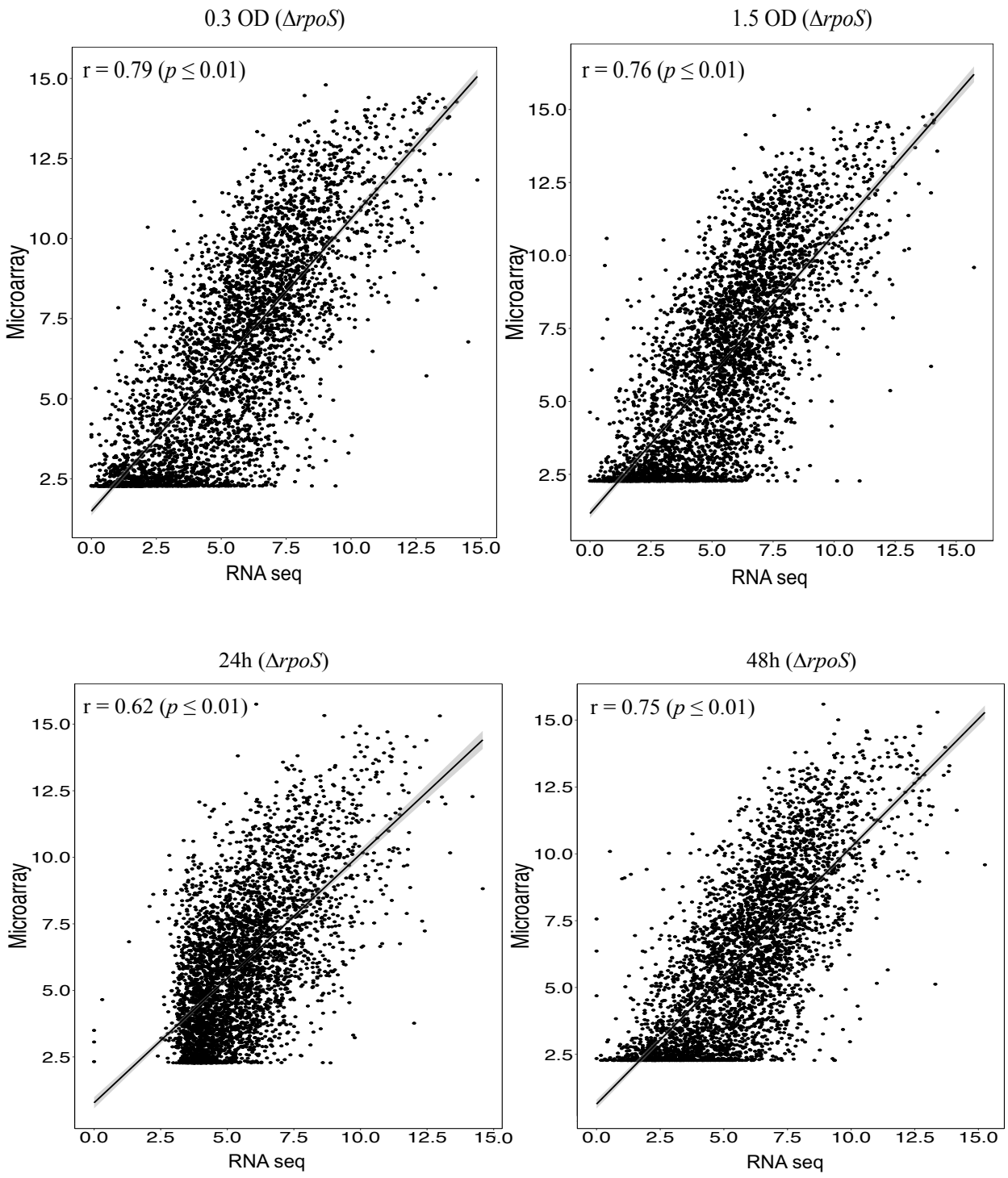


Figure 6: Pearson correlation analysis of RNA-seq and microarray data for Δ*rpoS*.
The relationship between the expression profiles generated by both platforms is depicted as a linear regression line. Pearson correlation coefficient represented by r value, and *p*-value shows the level of significance.

3.2 Validation of data with previous published work

To validate the current data, the high abundance transcripts during entry into early stationary phase (OD=1.5 relative to exponential phase) were compared with published data. RpoS controls about 10% of the genes during entry into stationary phase (Weber *et al.*, 2005). Additionally, RpoS-dependent genes such as glutamate-dependent acid resistance (*gadA, gadBE, hdeAB-yhiD, glsA-ybaT, slp-dctR*), components of Ni-Fe hydrogenase-1 (*hyaABCDEF*), the *csiD-lghO-gabDTP* encoding a carbon starvation protein, an L-2-hydroxyglutarate oxidase, and genes responsible for metabolism of γ-amino-butyric acid (GABA) (Patten *et al.*, 2004, Metzner *et al.*, 2004) were higher in abundance during the early stationary phase in both RNA-seq and microarray data. Moreover, the phosphate starvation induced gene *psiF*, stationary phase inducible aldehyde dehydrogenase, *aldB* and genes encoding nitrate reductase Z *(narZYWV*) were also among the most highly induced genes, which are also RpoS-dependent. *rmf* was also induced upon entry into stationary phase (Wada *et al.*, 1995). *rmf* encodes the ribosome modulation factor that inactivates 70S ribosome dimers by causing dimerization to 100S dimers. Furthermore, the down-regulation of flagellar biosynthesis genes during the early stationary phase (Patten *et al.*, 2004) was also confirmed in the current data. Altogether, the high abundance transcripts during the early stationary phase were mostly RpoS-dependent genes previously-identified in microarrays data (Patten *et al.*, 2004, Weber *et al.*, 2005), validating the current data. Moreover, the WT RNA-sequencing data was further validated by comparing with the previous generated data from our lab (unpublished data). The Pearson correlation coefficient was strong, suggesting good reproducibility of RNA-seq data (Appendix 11).

3.3 Overview of transcriptomic profile

While RpoS controls a large subset of genes during entry into stationary phase (about an hour-long transition between (OD600 = 0.3 to OD600 = 1.5), genes expressed after this transition remain largely uncharacterized. To address this, we performed global gene expression profiling at 24h and 48h of incubation using RNA-seq and microarray for MG1655 (WT) and *rpoS* mutant (Δ*rpoS*) strain. We observed no growth difference in both the strains and the generation time is also similar (Figure 7).

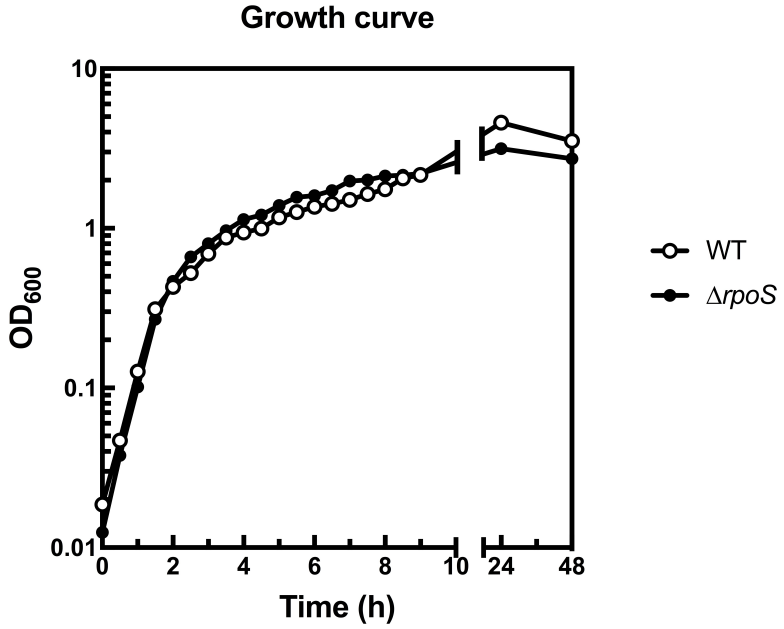


Figure 7: Growth curve of WT and Δ*rpoS* mutant strains in LB media.

Clustering of all growth phase samples

Clustering of replicates based on transcripts abundance was performed by Principal Component Analysis (PCA). If gene expression differences exist among the different growth phases, it should be expected that biological replicates of the same growth phase will cluster together in a principal component analysis. Consistently, the PCA for WT shows distinct clustering

of the three replicates for each growth phase, suggesting substantial change in gene expression between each growth phase but not much change in between replicates (Figure 8). Interestingly, for Δ*rpoS* strain 24h and 48h time-points replicates cluster closely together, suggesting fewer differences in in gene expression between the two time-points. However, 48h replicate 3 does not cluster with the other replicates, while the exponential (0.3 OD) and early stationary phase (1.5 OD) replicates are distinctly clustered together (Figure 9).

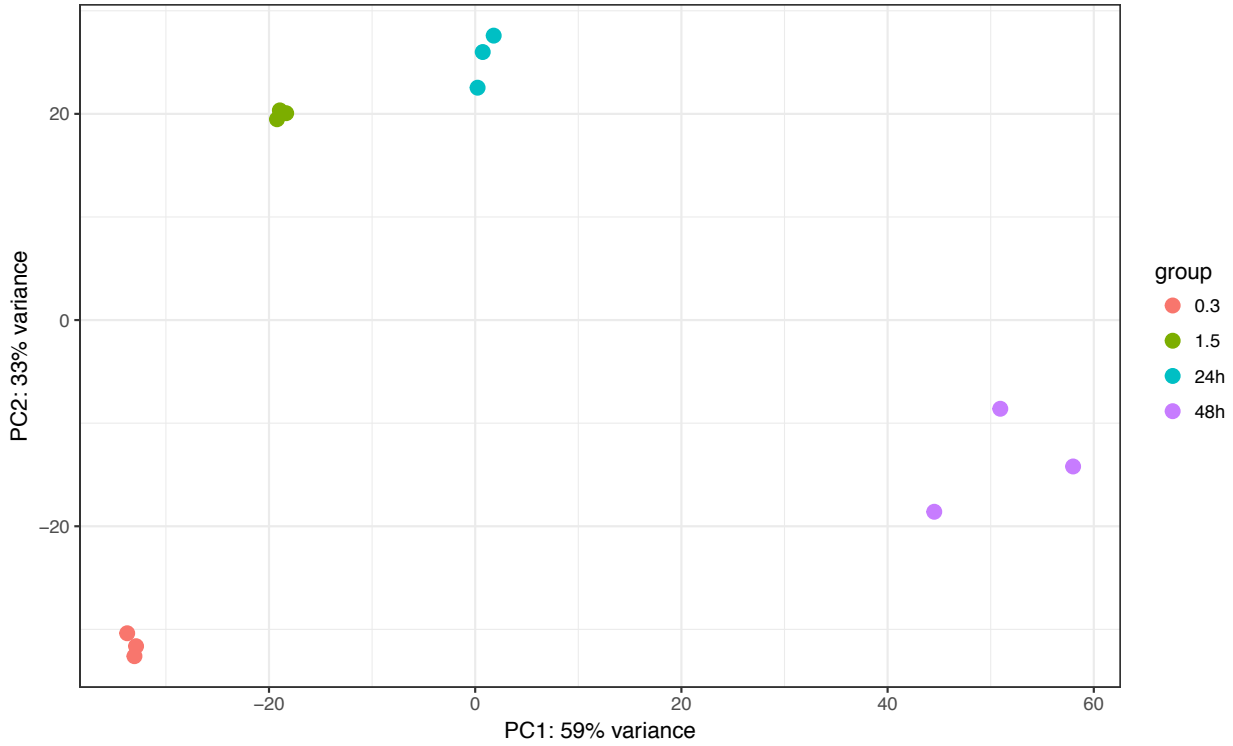


Figure 8: Principal component analysis (PCA) of the transcript abundance in WT.

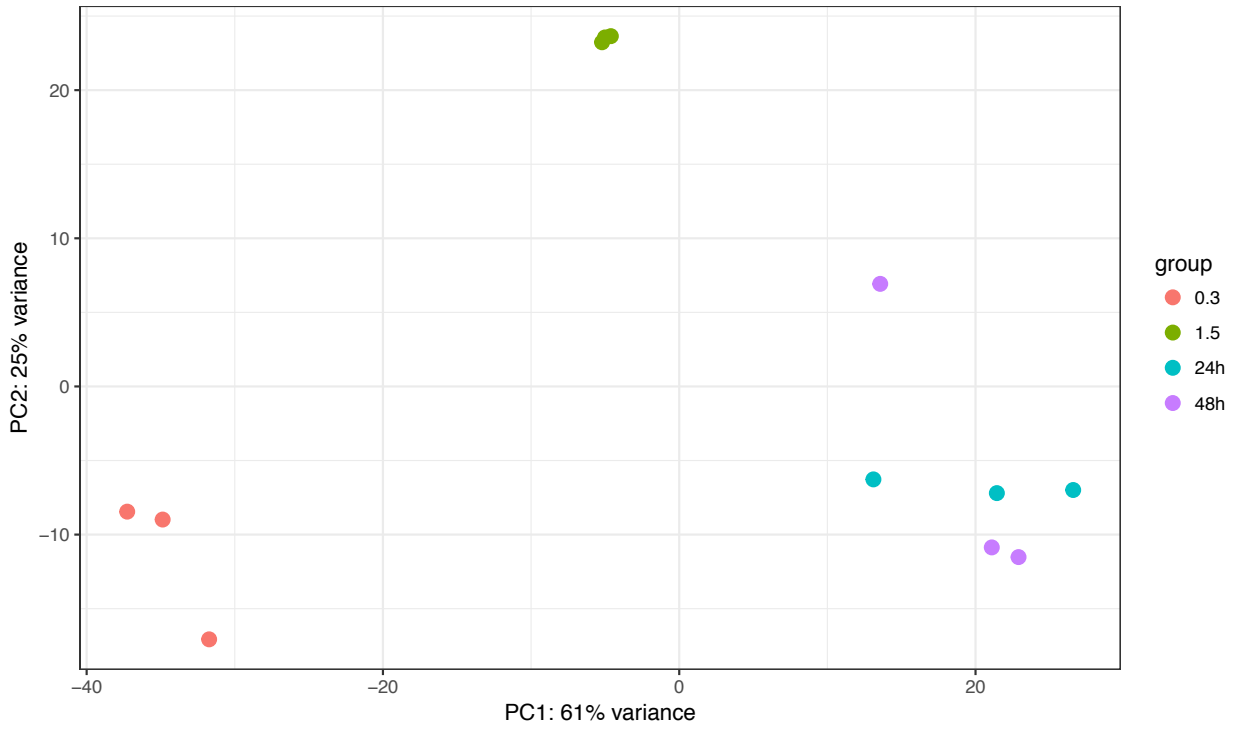


Figure 9: Principal component analysis (PCA) of the transcript abundance in Δ*rpoS*.

Differentially-expressed transcripts in RNA-seq and microarray data

Further we performed differential expression analysis to analyse the differentially-expressed transcripts in RNA-seq and microarray data. We are interested in identifying up-regulated transcripts that may be implicated in survival during prolonged-incubation phase. These were abundantly identified in RNA-seq technique (Table 4) and for functional enrichment analysis RNA-seq data was used. The overall summary for differentially-expressed transcripts in RNA-seq was also examined through MA plots (Appendix 4). Despite the threshold, there are few transcripts that are included in the discussion that showed a significantly lower fold-change than the selected threshold, considering their known functions in cell physiology and also it accounts for the biological processes of interest.

Table 4: Differentially-expressed transcripts in RNA-seq and microarray data.

	WT				Δ*rpoS*			
	24h relative to early stationary phase		48h relative to 24h		24h relative to early stationary phase		48h relative to 24h	
	RNA-seq	Micro array	RNA-seq	Micro array	RNA-seq	Micro array	RNA-seq	Micro array
High abundance transcripts	198	79	708	401	552	256	45	58
Low abundance transcripts	151	221	133	537	240	682	333	188

(Transcripts with a fold-change ≥ 4.0 with FDR adjusted $p \leq 0.05$ were considered as differentially-expressed.)

Which classes of genes are represented in highly abundant transcripts during prolonged-incubation phase?

To determine whether any functional category was over-represented in the group of differentially-expressed genes, functional enrichment analysis was performed on differentially up-regulated genes using the EcoCyc database (Karp *et al.*, 2014). Functional enrichment analysis determines if a gene set is statistically over-represented by genes within certain metabolic pathways, or by genes in certain Gene Ontology categories. The biological processes such as arginine catabolism, enterobactin biosynthesis, iron homeostasis, propionate catabolism and fatty acid catabolism were highly over-represented in both WT and Δ*rpoS* during initial prolonged-incubation (24h of incubation relative to early stationary phase) (Figure 10). This suggests that nutrient scavenging transcripts are in higher abundance and that their transcription does not require RpoS. At later prolonged-incubation phase (48h relative to 24h of incubation), the enriched biological processes in WT include cell adhesion, pilus organization and cell projection organization. Moreover, most highly induced transcripts encode proteins of unknown function, so it is difficult to predict their role during prolonged-incubation.

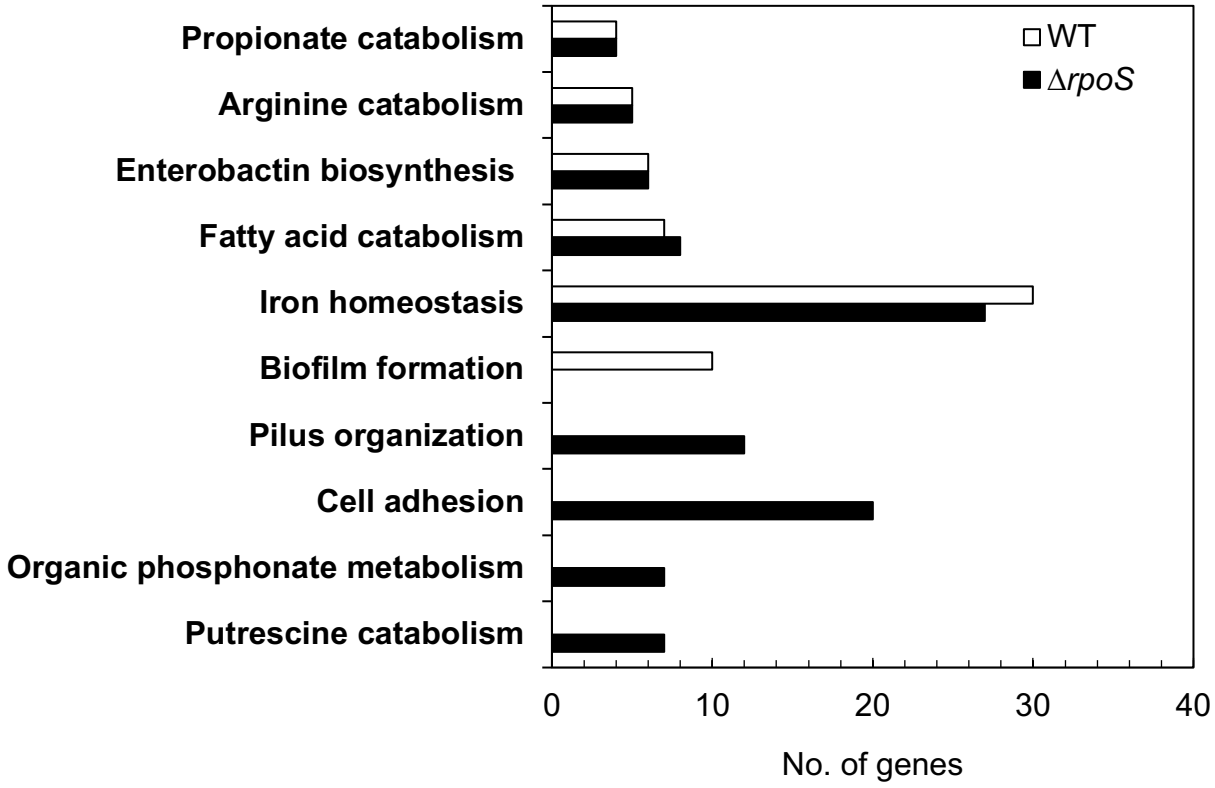


Figure 10: Over-represented biological categories for up-regulated genes during prolonged-incubation (24h) in WT and Δ*rpoS*.
Fisher's-exact statistical test with a Bonferroni correction method was used for significance testing (p-value ≤ 0.05).

Table 5: GO ontology classes over-represented within the gene set upregulated at 48h relative to 24h in WT and Δ*rpoS*.

WT	Δ*rpoS*
Pilus organization	Cellular amino acid metabolic process
Cell projection organization	Galactitol metabolic process
Cell adhesion	

Fisher's-exact statistical test with a Bonferroni correction method was used for significance testing (p-value ≤ 0.05).

Are the highly abundant transcripts during prolonged-incubation phase different from stationary phase regulon member?

Among the genes induced during entry into stationary phase, RpoS-dependent transcripts peaked during early stationary phase (OD600 = 1.5) and then declined during Prolonged-incubation phase. This was true for genes most highly dependent on RpoS for induction and its prototypical regulon members including *katE, osmY, dps*, *otsA (*Patten *et al.,* 2004) (Figure 11). This suggests that the transcripts induced during the prolonged-incubation phase are different from the stationary phase adaptation genes. Furthermore, the prolonged-incubation phase transcripts were compared with the other stress related conditions. The other related stress conditions include carbon starvation (Franchini & Egli, 2006), iron limitation (McHugh *et al.*, 2003), biofilm formation (Schembri *et al.*, 2003) and acidic condition (Kannan *et al.*, 2008). Those studies focussed on analyzing the differentially-expressed genes in particular stress conditions. To make the data comparable across other data, the genes with more than 2-fold expression (FDR adjusted p-value ≤ 0.05) during-prolonged incubation were selected as other data have the same criteria. The genes expressed during prolonged-incubation are unique since the overlap is mostly observed with the genes expressed during iron-limited and carbon starvation condition. However, the biofilm study showed the least overlap with the prolonged-incubation phase (Figure 12), even though the biofilm-related transcripts were higher in abundance during prolonged-incubation. This suggests that biofilm-related genes expressed during prolonged-incubation in planktonic cultures are different from the genes expressed during biofilm formation.

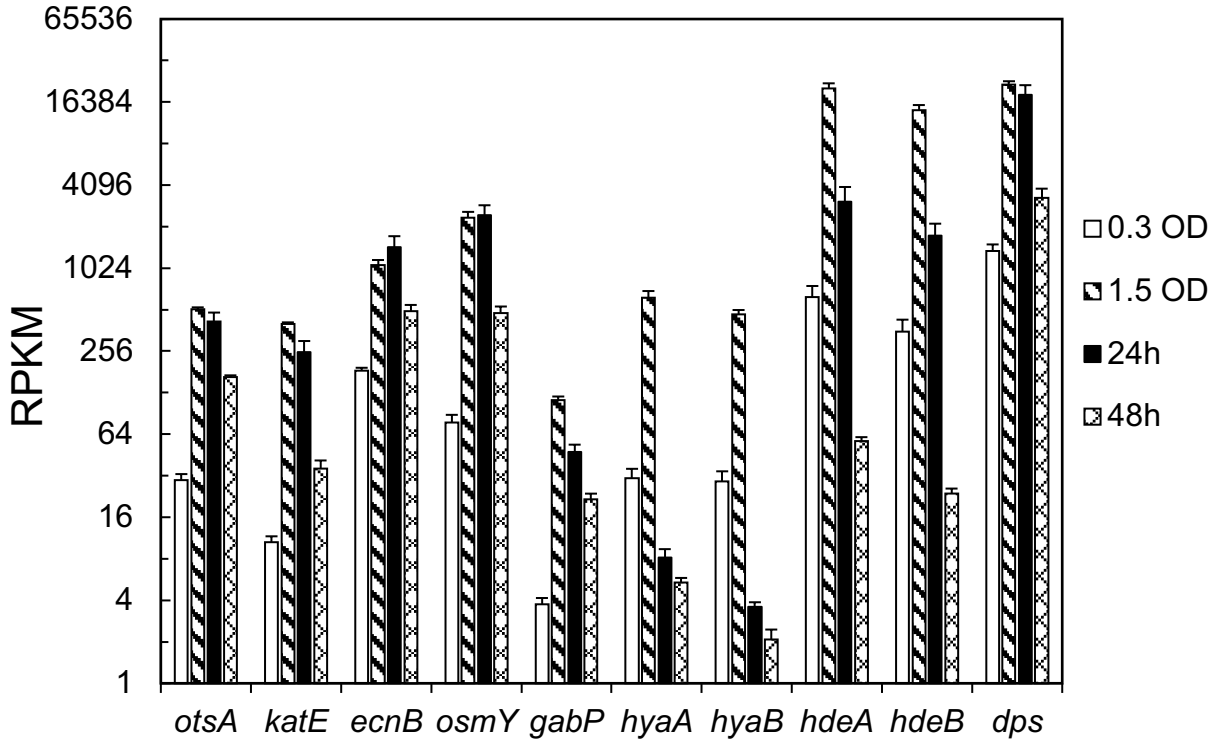


Figure 11: RpoS-dependent transcripts decreased in abundance during the prolonged-incubation phase.
Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

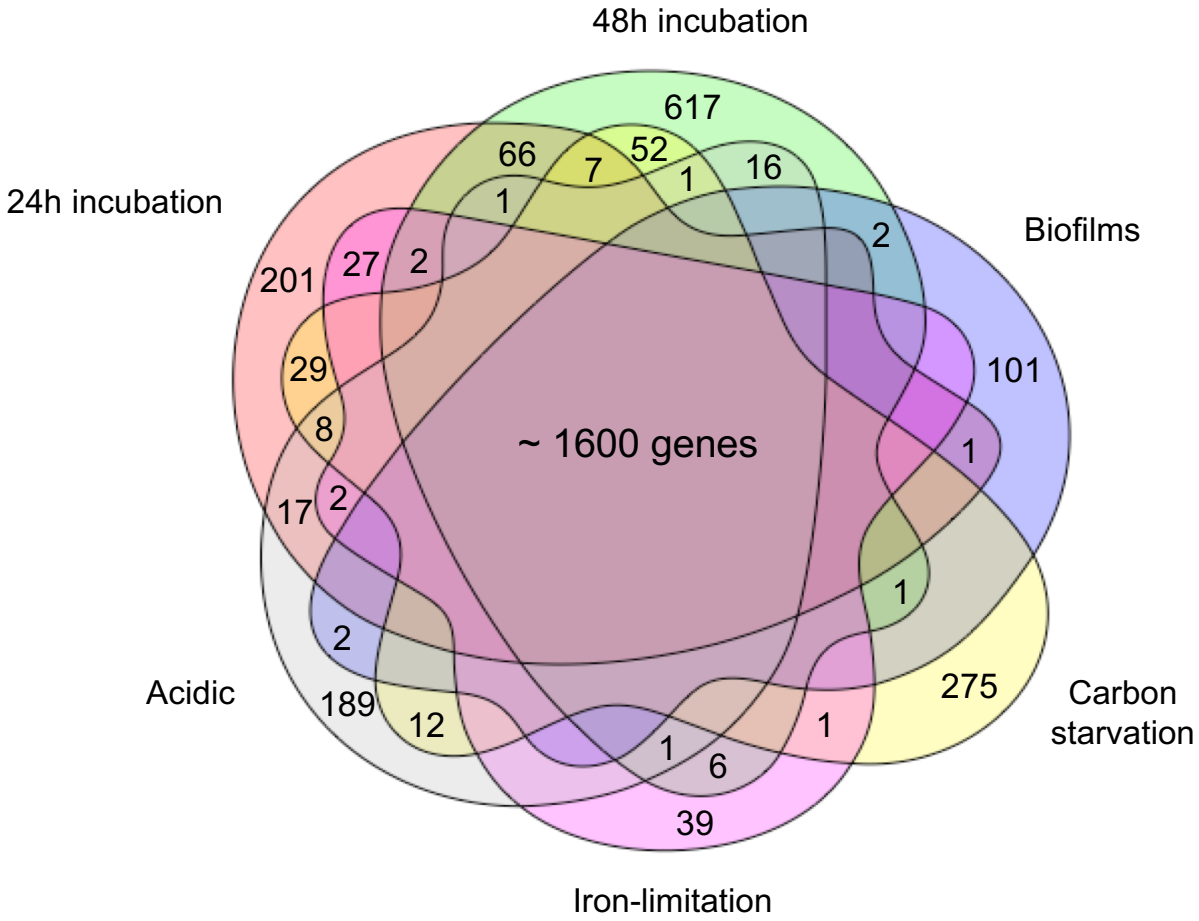


Figure 12: Specific genes expressed during prolonged-incubation in compare to other related stress conditions.
Carbon starvation (Franchini & Egli, 2006), Iron-limitation (McHugh et al., 2003), Biofilms (Schembri et al., 2003) and Acidic (Kannan et al., 2008). For the prolonged-incubation phase, the highly abundant transcripts with more than 2-fold change in expression were included, to make the data comparable with other conditions.

Are highly-abundant transcripts during prolonged-incubation are RpoS-independent or dependent?

To determine the RpoS-independent or dependent genes, the differentially-expressed transcripts in WT and Δ*rpoS* during prolonged-incubation phase were compared (Figure 13). For instance, at 24h of incubation relative to early stationary phase, the transcripts showed high abundance in WT and also in Δ*rpoS* mutant were considered RpoS-independent. Alternatively, if the transcripts showed high abundance in WT and low abundance in Δ*rpoS* were considered RpoS-dependent. The data indicates that both RpoS-dependent and independent genes are expressed during prolonged-incubation, where during initial prolonged-incubation (24h relative to early stationary phase) RpoS-independent transcripts are high and during latter prolonged-incubation (48h relative to 24h) RpoS-dependent transcripts are high. The following is a discussion of genes of interest and their physiological function with relevant references and information from the EcoCyc database (a complete list of the genes and their associated fold-changes during prolonged-incubation can be found in the Appendix 5 and 6).

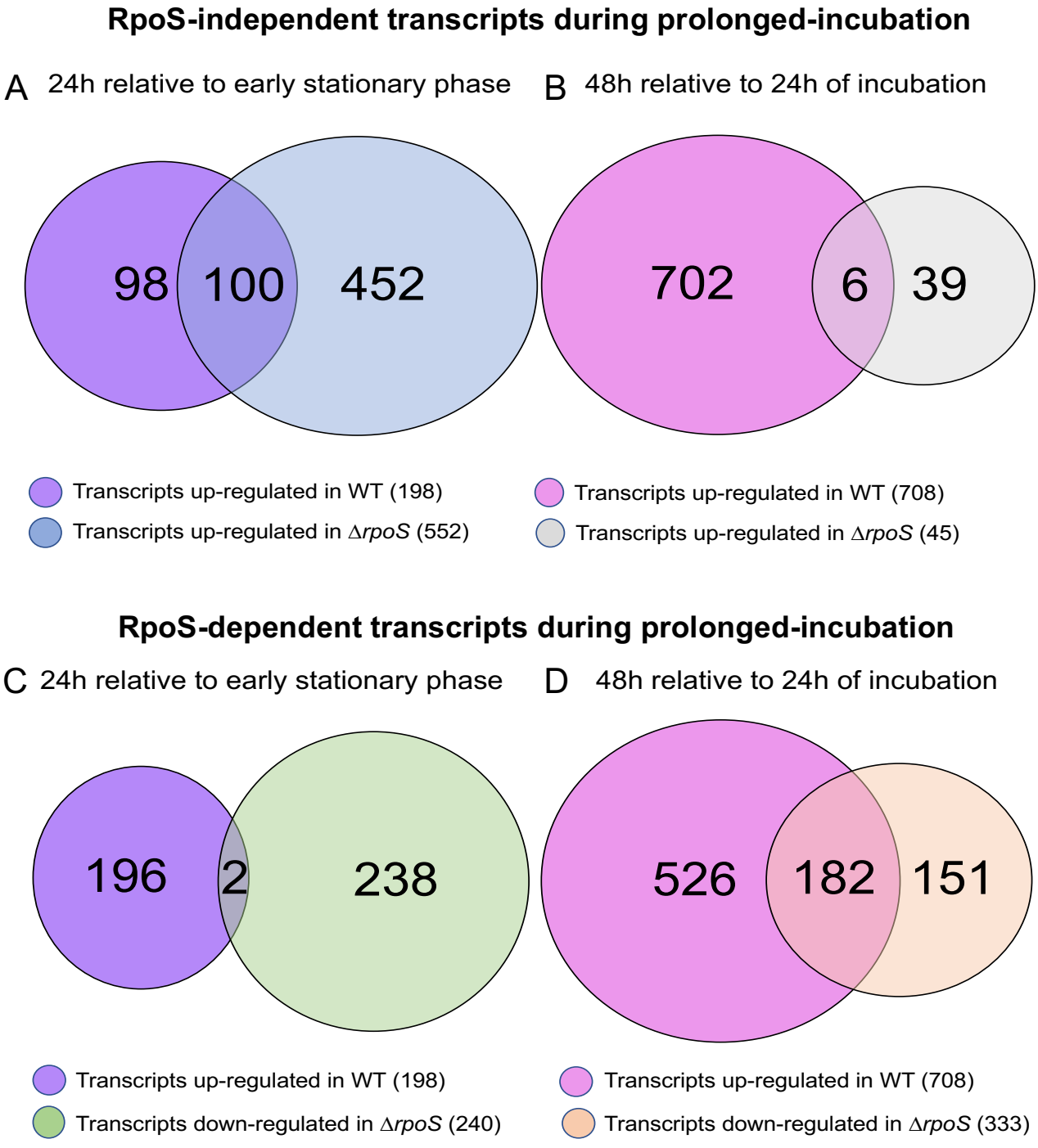


Figure 13: Comparison of differentially-expressed transcripts in WT and Δ*rpoS* to determine RpoS-independent or dependent transcripts during prolonged-incubation.
The high abundant transcripts in WT compared with high abundant transcripts in Δ*rpoS* during prolonged-incubation (A and B) considered as RpoS-independent. The high abundant transcripts in WT compared with the low abundant transcripts in Δ*rpoS* during prolonged-incubation (Band D) considered as RpoS-dependent. Values in brackets represent the total number of significant transcripts in that particular condition (fold-change ≥ 4 and FDR adjusted $p \leq 0.05$).

3.4 RpoS-independent transcripts

3.4.1 Iron acquisition genes

Transcripts that are repressed by ferric uptake regulator, Fur, in normal growth condition were among the most highly induced transcripts during prolonged-incubation (24h of incubation relative to early stationary phase). The transcripts are responsible for maintaining iron homeostasis within the cells (Bagg & Neilands, 1987). Moreover, in the current study, these transcripts were expressed as RpoS independent during prolonged-incubation. The expression of iron acquisition genes during prolonged-incubation may be due to iron limitation/oxidation during prolonged-incubation. Fur, acts as a repressor, along with iron as a co-repressor $[Fur\text{-}Fe^{2+}]^2$ of the genes encoding proteins for iron acquisition and siderophore-mediated iron transport (Bagg & Neilands, 1987). *fur* transcription is autoregulated (repressed) by $[Fur\text{-}Fe^{2+}]^2$ and activated by cAMP-CRP, linking iron metabolism to carbon metabolism within the cell (Escolar *et al.*, 1999). Additionally, *fur* is activated by the *oxyR* and *soxRS* oxidative stress regulators and *fur* represses *soxS* (Zheng *et al*., 1999). It is unclear in our study as to what signal triggers the expression of Fur regulon. An iron-limitation may be a strong inducing signal that regulates gene expression in prolonged-incubation phase cultures, however, this hypothesis remains to be tested.

Prolonged-incubation is characterized by a low-nutrient environment in which low levels of soluble iron can limit growth. Insoluble iron may be present in the environment, but it can be utilized by *E. coli* unless it is bound by iron-chelating molecules. To survive and acquire the iron necessary for growth in iron-limiting conditions, *E. coli* secretes siderophore (enterobactin). Enterobactin is a prototypical catecholate siderophore that has a high affinity for iron (Winkelmann, 2002). Enterobactin solubilizes the extracellular iron, by reduction or chelation, followed by internalization with the specific transporter proteins (Braun, 2003). Consistent with

that, the transcripts for the genes of the enterobactin biosynthesis pathway along with its transporter protein-coding transcripts (Figure 14) were in higher abundance during prolonged-incubation.

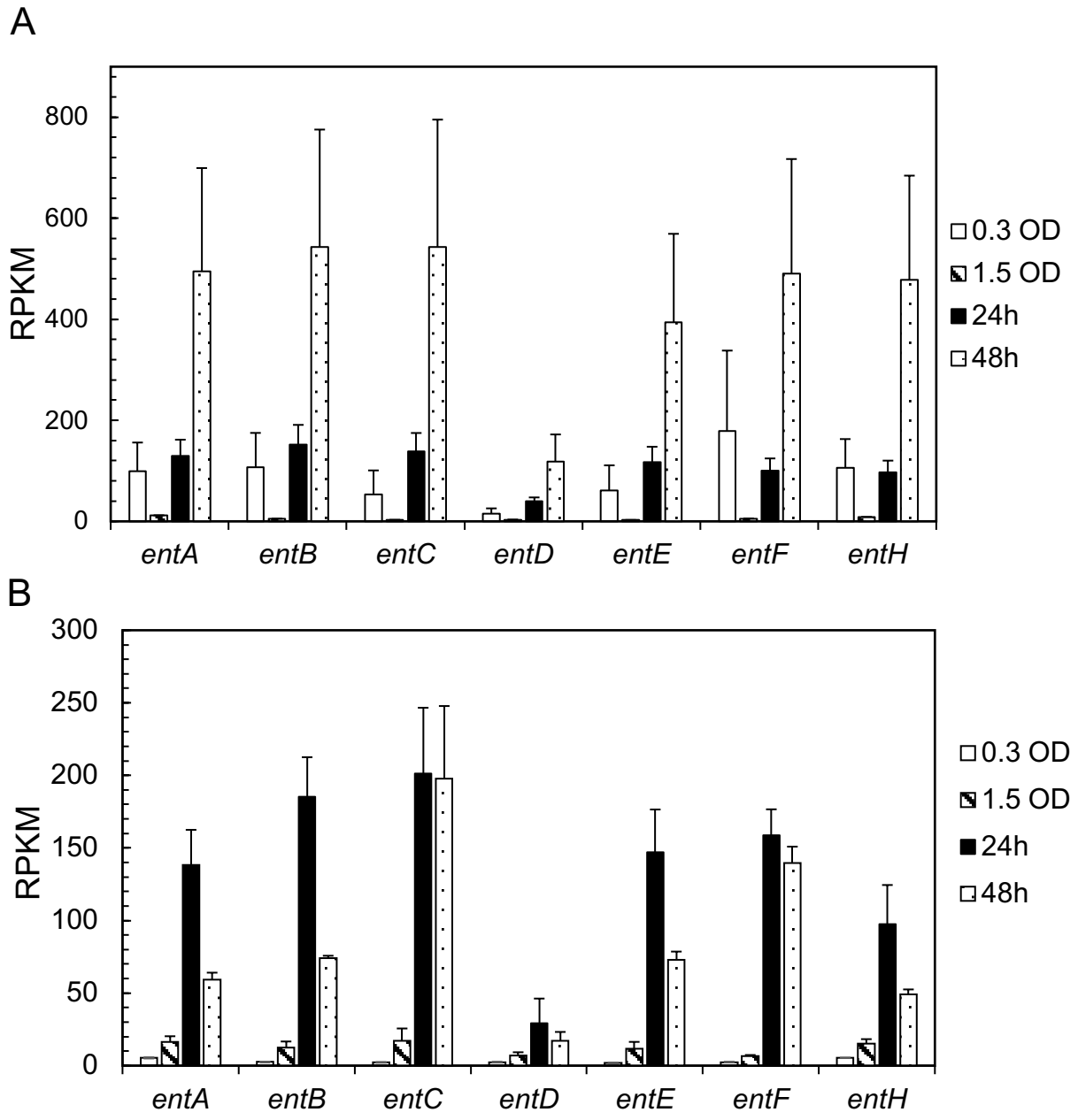


Figure 14: High transcript abundance of the Enterobactin biosynthesis pathway genes during prolonged-incubation.
The transcript abundance in WT (A) and Δ*rpoS* (B). Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

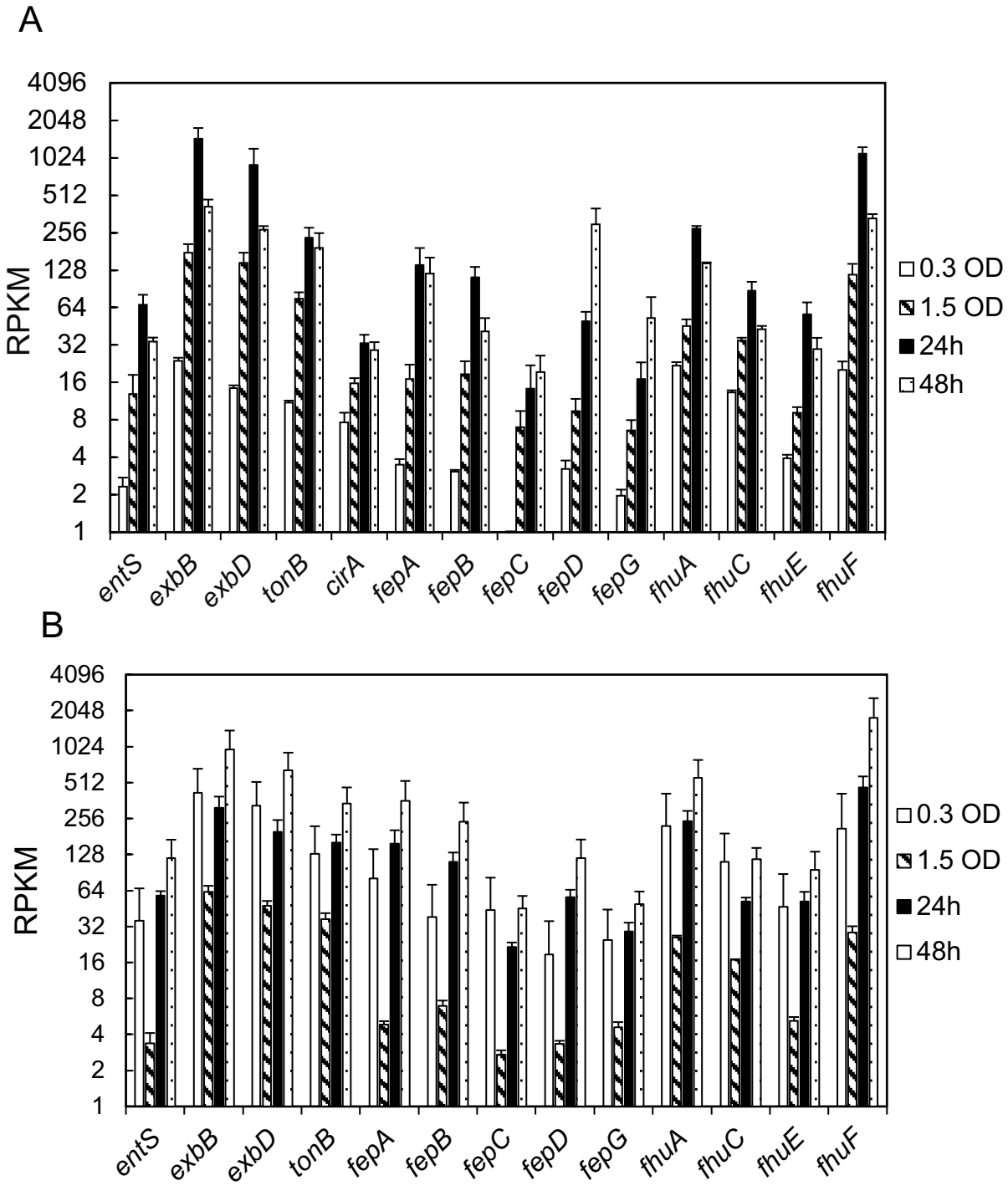


Figure 15: High transcript abundance of the iron transporter genes during prolonged-incubation.

The transcript abundance in WT (A) and in Δ*rpoS* (B). Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

Another adaptive mechanism mediated by *E. coli* to compensate for iron deficiency is to import manganese ion to serves as a substitute for iron (Martin *et al.*, 2015). Manganese is less readily oxidized than iron. Substitution of Fe^{2+} with Mn^{2+} preserves the function of non-redox enzymes that would otherwise utilize Fe^{2+} as a cofactor and become inactivated when it is oxidized. MntH is a high-affinity H^{+}/Mn^{2+} symporter that imports Mn^{2+} into the cell and is up-regulated by Fur mediated de-repression and through OxyR during oxidative stress. *mntS* encodes a protein that increases the intracellular Mn^{2+} concentrations by interfering with the action of MntP, a Mn^{2+} exporter, or by acting as a Mn^{2+} chaperone (Martin *et al.*, 2015). Therefore, it was expected that genes responsible for Mn^{2+} acquisition and retention would be up-regulated in prolonged stationary phase cultures. This hypothesis was supported by increased expression of *mntH, mntS* and *nrdHIEF* operon (Figure 16). The *nrdHIEF* operon encoding the Mn^{2+}-dependent ribonucleotide reductase system. This provides deoxyribonucleotide precursors for DNA synthesis using Mn^{2+} under the iron starvation condition. Fur regulates the expression of *nrd* operon (Seo *et al.*, 2014). Moreover, under oxidative stress, the expression of this operon is increased (Monje-Casas *et al.*, 2001).

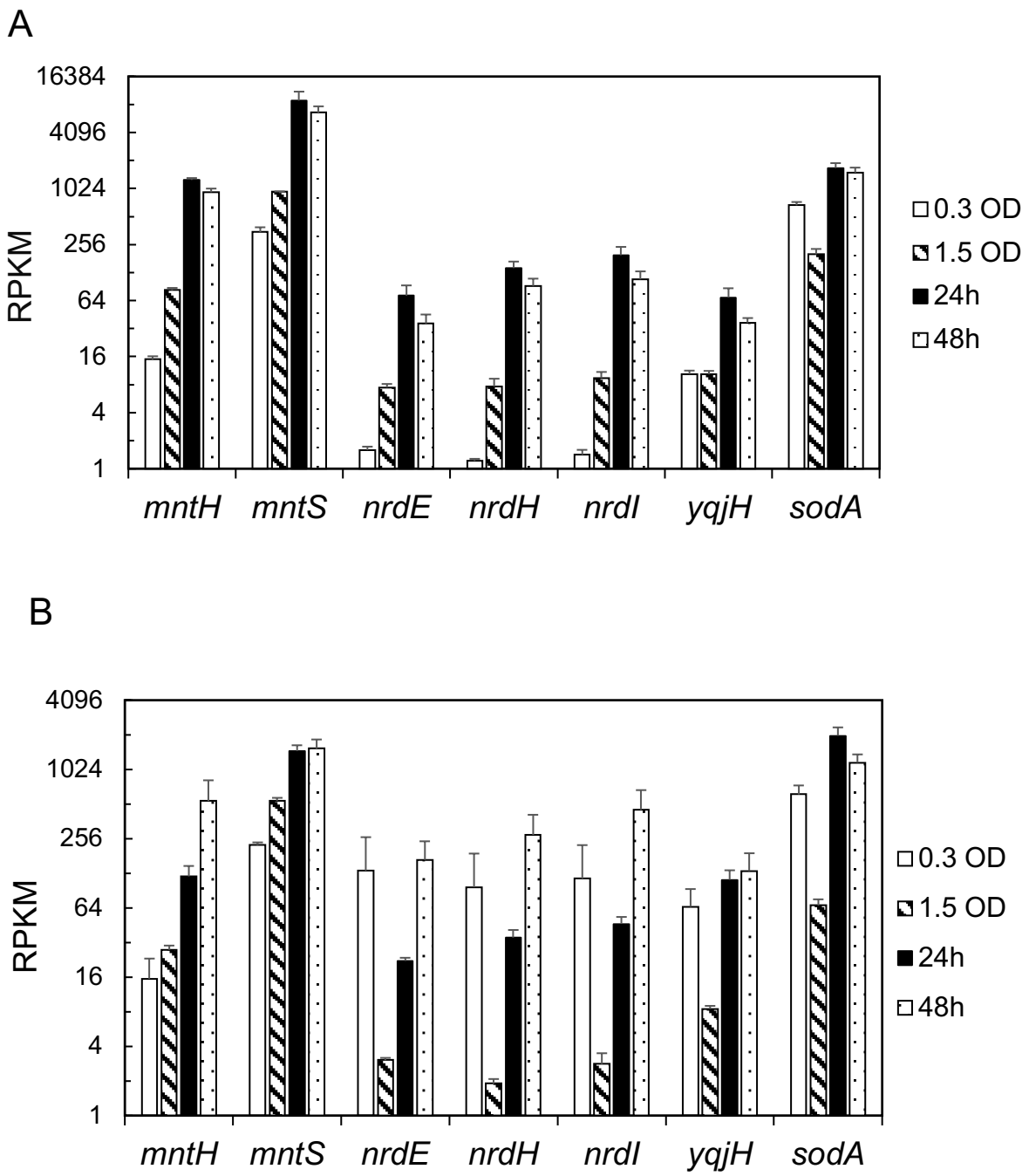


Figure 16: Mangenese transporter and Mn^{2+} dependent transcripts were in higher abundance during prolonged-incubation.
The transcript abundance in WT (A) and in Δ*rpoS* (B). Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

The iron-limitation may also be caused by oxidation of enzyme-bound iron during oxidative stress. Oxidative stress or iron-limitation can also cause perturbation in the levels of 2Fe-2S clusters required for the function of many enzymes. IcsR is a 2Fe-2S containing transcription factor that senses the intracellular levels of 2Fe-2S clusters and derepresses the transcription of the Isc (*iscS*, *iscU*, and *iscA)* and Suf systems (*sufABCDS*), which encodes enzymes responsible for the 2Fe-2S cluster. (Giel *et al.*, 2006). Moreover, a double mutant of both operon exhibits synthetic lethality, indicating that these systems are redundant and that iron-sulfur cluster assembly is essential for viability (Takahashi & Tokumoto, 2002). Consistent with the iron limitation/oxidation hypothesis, we observed an increase in expression of *isc* and *suf* operons during prolonged-incubation (Figure 17).

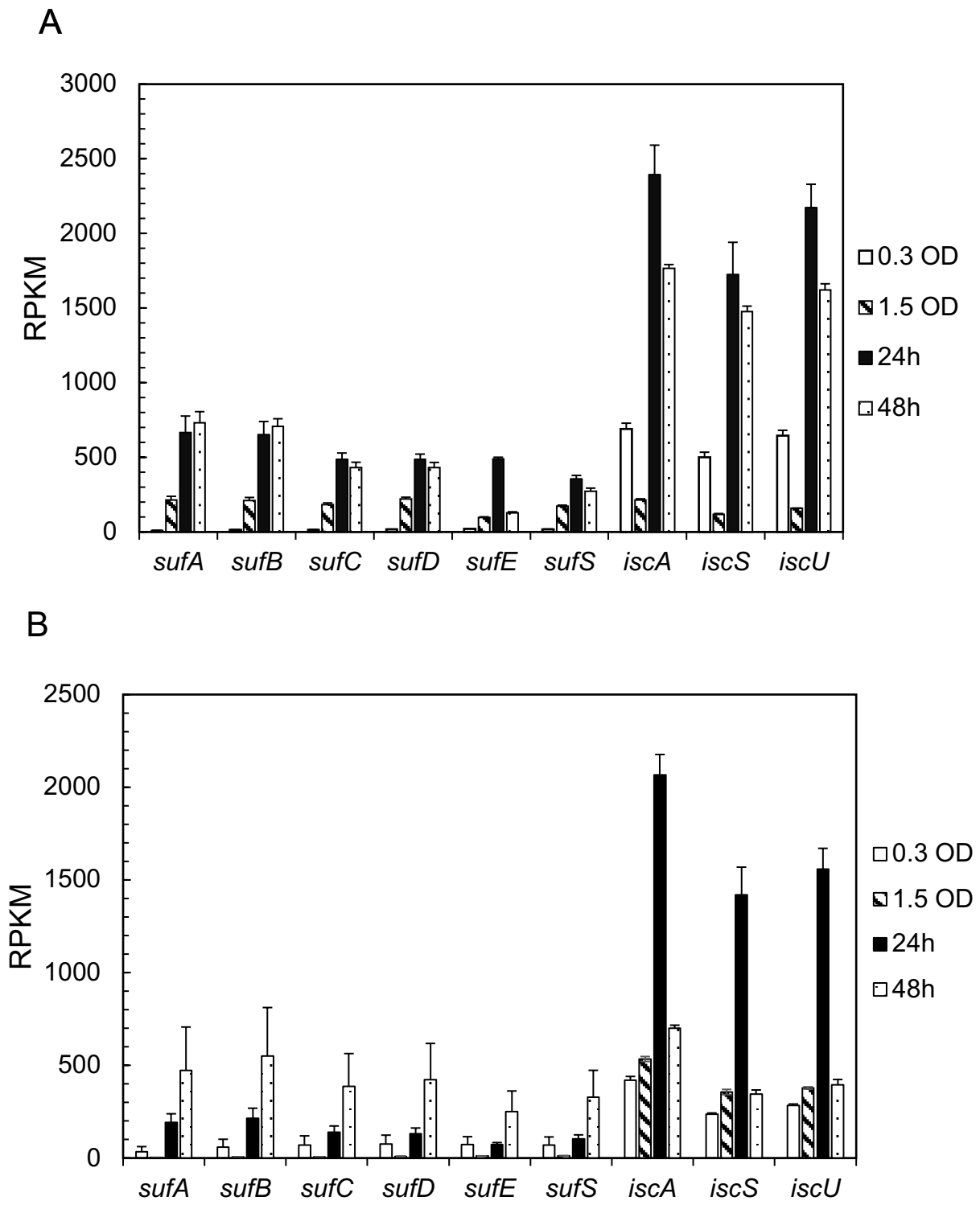


Figure 17: FeS cluster assembly transcripts were in high abundance during prolonged-incubation phase.
The transcript abundance in WT (A) and in Δ*rpoS* (B). Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

Altogether, these findings reinforce the idea that cells in prolonged-incubation phase may experience iron limitation perhaps as a result of limiting amounts available in LB or loss of iron due to oxidative stress. Moreover, consistent with the hypothesis that the many growth phase regulated functions in *E. coli* do not require RpoS for expression and the other regulatory mechanism, in addition to RpoS may control gene expression during prolonged-incubation.

3.4.2 Degradation processes

Along with iron limitation, the transcripts for degradation enzymes responsible for catabolic processes are also in higher abundance during prolonged-incubation. The low nutrient level in prolonged-incubation phase cultures may act as a signal for their expression. *E. coli* cells undergo alteration in carbon nutrition. As the cell utilizes glucose, acetate is produced as a product under aerobic respiration and that is exported from the cell through the phosphotransacetylase-acetate kinase pathway (Kumari *et al.*, 2000). Upon entry into stationary phase, the acetate is up-taken by the cell and utilizes acetate as a carbon source (Akesson *et al.*, 1999). Consistently, the transcripts for acetyl-CoA synthetase enzyme (*acs*) along with its transporter protein-coding transcripts (*actP*) were in higher abundance during prolonged-incubation. Acetyl-CoA synthetase enzyme converts acetate to acetyl-CoA known as ACS pathway and fed into the tricarboxylic acid pathway (TCA) and glyoxylate shunt (Kumari *et al.*, 2000). The ACS pathway function in an anabolic role, scavenging acetate present in the extracellular medium. Induction of *acs* expression functions as the metabolic switch activating this pathway (Valgepea *et al.*, 2010). Acetyl-CoA synthetase (*acs*) is reported to be under the control of cAMP, Fnr and the flux of carbon through the acetate pathway (Renilla *et al.*, 2012).

When acetate is used up, the cell starts to utilize amino acids as carbon and nitrogen sources during the stationary phase. The cells consume easy to utilize amino acids (L-serine, L-aspartate,

L-tryptophan, L-glutamate, L-glycine, and L-alanine) until they are depleted, then switch to the harder to utilize amino acids (L-arginine, L-glutamine, L-asparagine, L-cysteine and L-lysine). *E. coli* in tryptone broth culture sequentially catabolizes these amino acids and in LB broth as well since a diauxic lag is observed and it may also follow the same order (Sezonov *et al.*, 2007). Consistent with this, the transcripts for genes coding enzymes of L-arginine degradation (AST II) pathway (Figure 18) were highly abundant during prolonged-incubation. The arginine catabolic process, in which arginine is converted to succinate and glutamate, yields two molecules of ammonia. As ammonia is a good source of nitrogen, this can satisfy the total nitrogen requirement of *E. coli* (Schneider *et al.*, 1998). Furthermore, nitrogen limitation induces AST II pathway enzymes and elevates the transcripts level. In *E. coli,* the *astCADBE* operon contains two promoters, an Ntr-dependent promoter that requires σ^{54} (RpoN) and a phosphorylated NtrC. However, in the stationary phase, the transcription initiates from a σ^{S} (RpoS)-dependent promoter that is 5 bases downstream of the Ntr promoter (Kiupakis & Reitzer, 2002). In the current study, the *ast* operon expression does not require RpoS during prolonged-incubation in rich media.

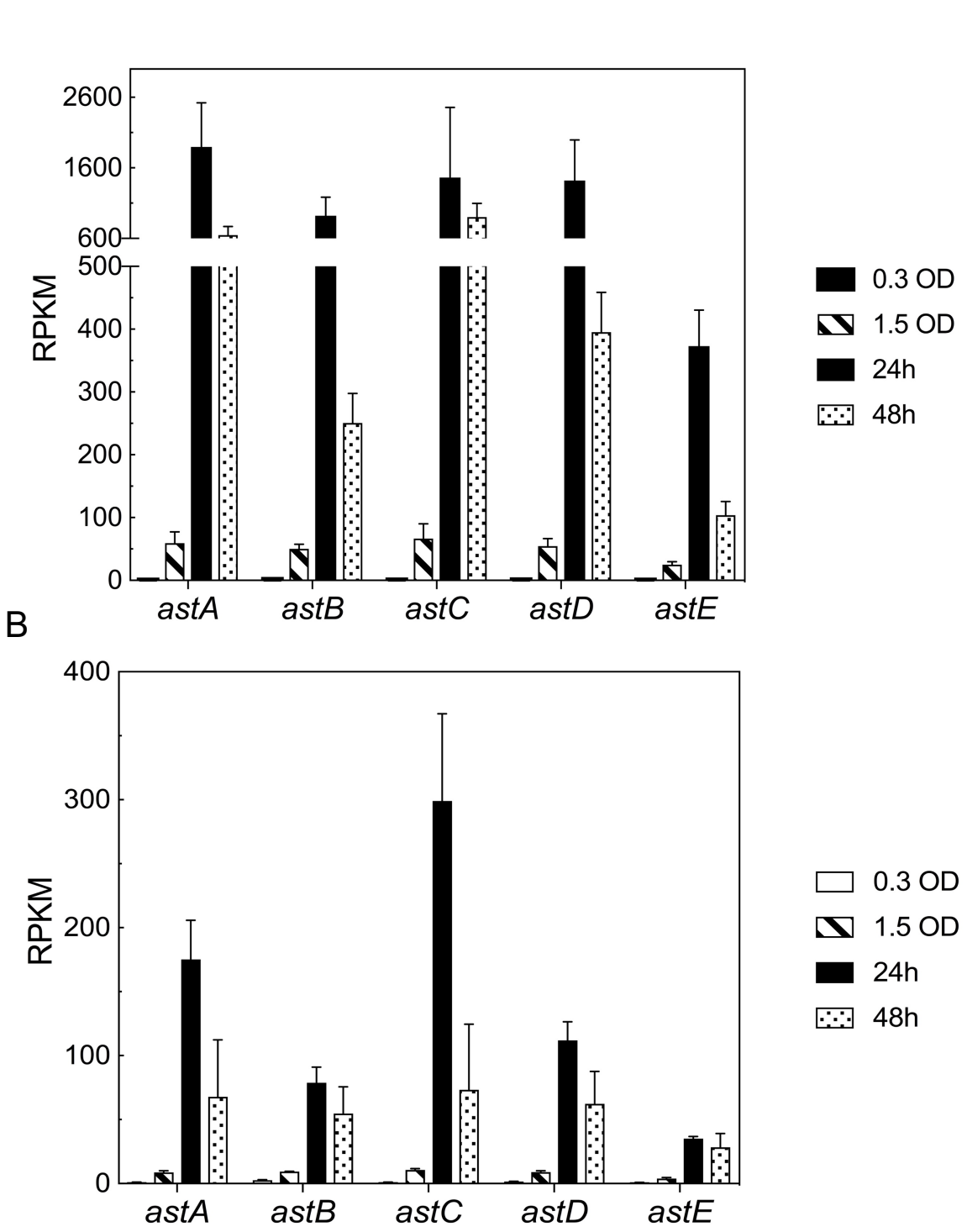


Figure 18: High transcript abundance of the L-arginine degradation II (AST pathway) pathway genes during prolonged-incubation.
The transcript abundance in WT (A) and Δ*rpoS* (B). Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

Moreover, the genes for carboxylate degradation that are under the control of RpoN were strongly expressed during prolonged-incubation. This includes transcripts for the genes encoding enzymes for the 2-methylcitrate pathway (propionate degradation) (Figure 19) were abundant and expressed in a RpoS-independent manner. The *prpBCDE* operon codes for proteins needed for catabolism of propionate and *prpR* act as a regulatory protein for the operon. In addition to RpoN, the genes belonging to the propionate metabolism are also regulated by cAMP receptor protein (CRP) and integration host factor (IHF) (Lee *et al.*, 2005). However, the *prp* operon is strongly under RpoS control during stationary phase in minimal (Dong & Schellhorn, 2009a) and also in glucose-limited media (Franchini *et al.*, 2015). On the other hand, the current finding suggests that during prolonged-incubation the transcription of *prp operon* does not require RpoS in rich media.

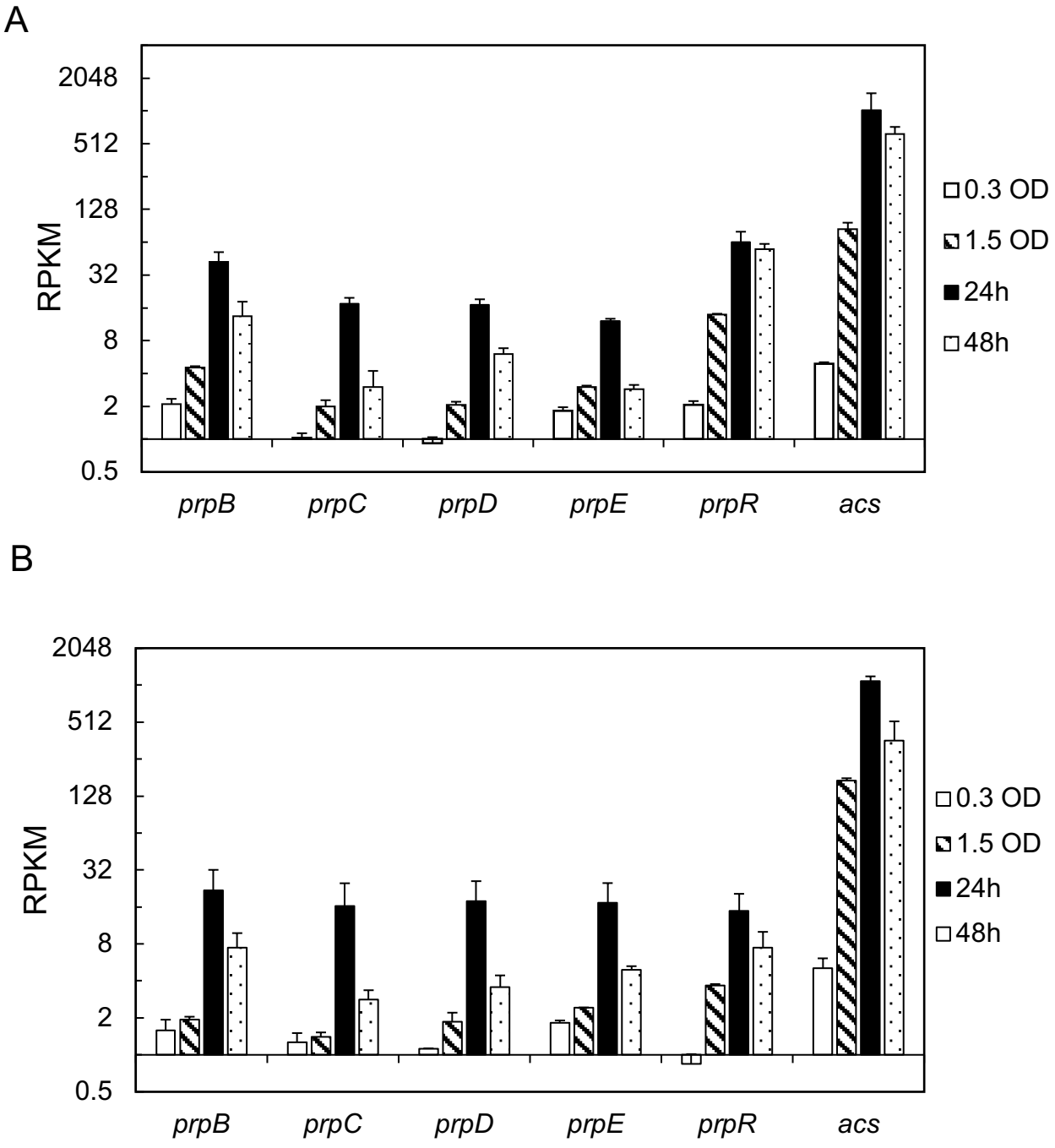


Figure 19: High transcript abundance of the Propionate degradation (2-methylcitrate cycle I) pathway genes during prolonged-incubation.
The transcript abundance in WT (A) and Δ*rpoS* (B). Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

Fatty acids degradation *fad* genes (Figure 20) were also preferentially expressed during prolonged-incubation and also in a RpoS-independent fashion. *E. coli* can use fatty acids with diverse chain lengths as its sole carbon and energy sources. After uptake, fatty acids can either be degraded through the ß-oxidation pathway or used as precursors for membrane phospholipid biosynthesis. The degradation pathway enzymes are encoded by the *fad* regulon, which are responsible for the transport (*fadD, fadL*) and activation of long-chain fatty acids (*fadE)*, and their ß-oxidative (*fadBA)* cleavage into acetyl-CoAs. The expression of genes encoding the fatty acid oxidative enzymes is negatively controlled by fatty acids-specific FadR regulator. The ArcAB system strongly represses the expression of the 3-hydroxyacyl-coenzyme A (CoA) dehydrogenase encoded by the *fadB* gene and weakly represses acyl-CoA dehydrogenase activity encoded by *fadE* gene (Cho *et al.*, 2006). The mechanism(s) of repression of these genes by the ArcAB system have not yet been explored. The specific regulatory mechanism exerted by the FadR transcriptional factor plays a dual role in fatty acid metabolism. FadR specifically represses the transcription of *fad* regulon and activates the unsaturated fatty acids biosynthesis. During entry into stationary phase, FadR derepresses *fad* genes, which suggests that the FadR regulation may be responsible for providing the growth-arrested cells with endogenous carbon and energy from membrane-derived fatty acids (Farewell *et al.*, 1996). Furthermore, the levels of the alarmone ppGpp increases during stationary phase, which leads to a decrease in fatty acids biosynthesis and there is an accumulation of fatty acids biosynthesis product, long-chain acyl-carrier protein (long-chain acyl-ACP). Concurrently, the long-chain acyl-ACP converts to long-chain fatty acids acyl-CoA (LCACoA). Elevated levels of LCACoA causes the inhibition of FadR-dependent DNA binding which allows induction of *fad* genes (DiRusso & Nyström, 1998). In addition, if the growth of cells is arrested due to lack of carbon source an increase in cAMP-CRP will further amplify the

induction of *fad* genes (Feng & Cronan, 2009). In our study, however, the signal triggers the induction *fad* regulon during prolonged-incubation is unclear. Altogether, the data suggest that during prolonged-incubation maybe the cells sense the nutrient starvation condition and activates the different catabolic pathways to manage their nutrient and energy requirements.

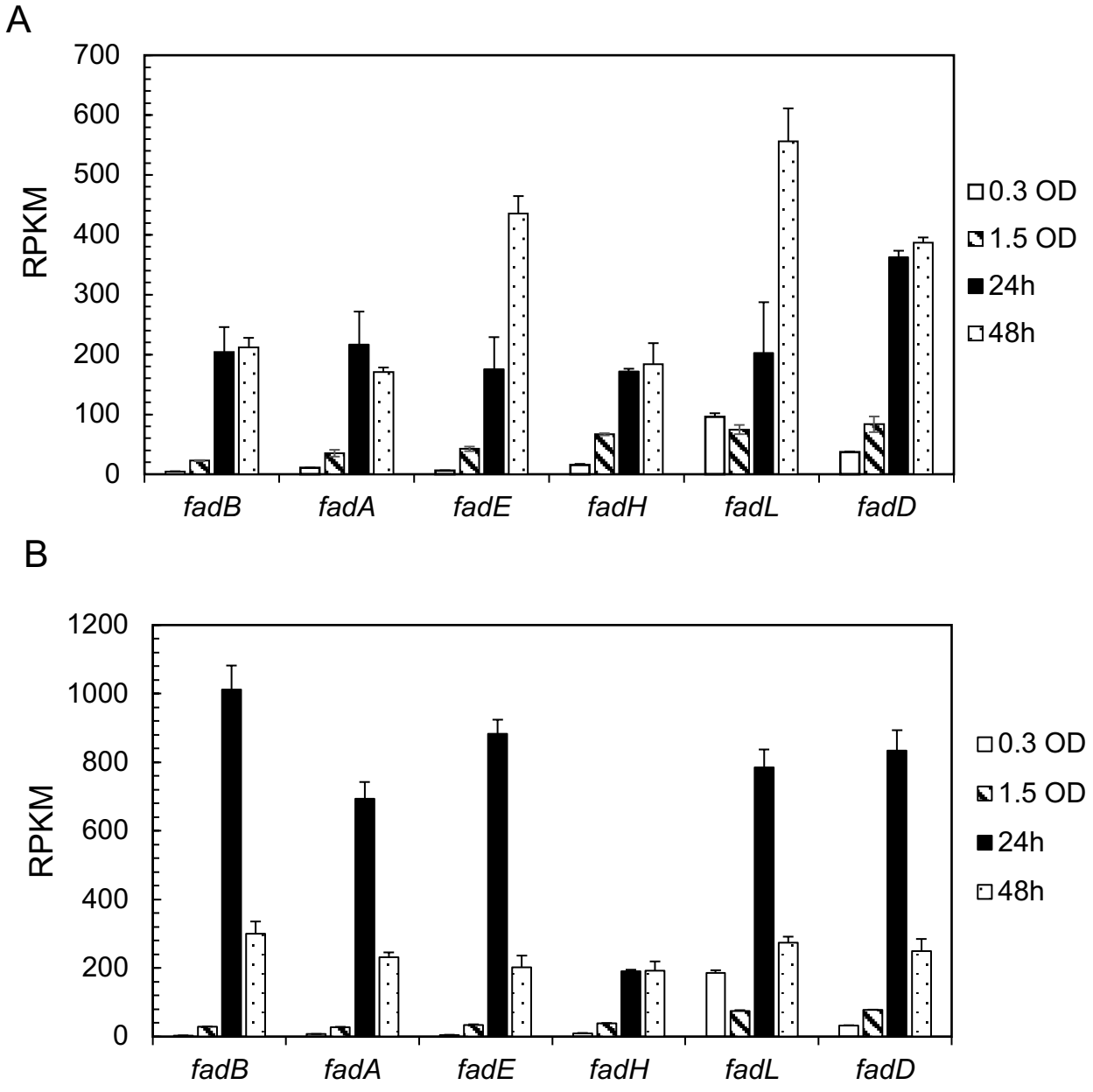


Figure 20: High transcript abundance of the fatty acid degradation genes were during prolonged-incubation.
The transcript abundance in WT (B) and Δ*rpoS* (C). Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

3.5. RpoS-dependent transcripts

3.5.1 Adhesion and Fimbriae genes

Seven cryptic but functional chaperone *fim*-like operons were identified in *E. coli* (Badouraly *et al.*, 2010). These genes encode functional fimbriae adhesins, transported by the chaperone-usher (CU) secretion system that promotes adhesion to abiotic and/or biotic surfaces. The operons are *sfmACDHF, ycbQRSTUVF, yraHIJK*, *yadNVhtrEyadMLKC, yehABCD, ybgOPQD,* and *yfcOPQRSTUV*. The expression of these operons are not observed during normal laboratory growth but constitutively expressed, *E. coli* form fimbriae like structures (Badouraly *et al.*, 2010). Interestingly, in WT, these operons were highly up-regulated during prolonged-incubation (48h relative to 24h of incubation) (Appendix 9). *ydeQRST*, *gltFyhcADEF* operons and *ygiL* and *yagY* transcripts which possess strong sequence and organization homologies to the Type 1 fimbriae *fim* operon (Nuccio & Bäumler, 2007) were also up-regulated. The genes mention in Table 6 showed RpoS-dependent expression during prolonged-incubation. The constitutive expression of *yad*, *ycb* and *yeh* operon promotes biofilm formation on different abiotic surfaces in the absences of Type 1 fimbriae. Furthermore, the *yfc*, *yra* and *sfm* also promote bacterial adherence on eukaryotic cells. (Badouraly *et al.*, 2010).

Furthermore, other adhesion-like protein-coding genes were expressed during prolonged-incubation, include the adhesion-like autotransporter gene, *yejO* that shows sequence similarity to the surface expressed antigen 43 (Henderson & Owen, 1999). Antigen 43 is an autotransporter adhesin which contributes to either colonization or biofilm maturation. This protein contributes to cell-to-cell adhesions after the initial attachment to an abiotic surface. (Kjaergaard K, 2000). *ypjA* is another adhesion-like autotransporter coding gene that up-regulated during 48h relative to 24h of incubation. However, deletion of the *ypjA* gene has no significant effect on adhesion to solid surfaces (Roux *et al.*, 2005). The putative porin-domain gene *eaeH*, whose protein sequence is

highly similar to the conserved protein EaeH of enterotoxigenic *E. coli* strains also induced in expression. In ETEC strains, this protein encodes adhesion which contributes to bacterial adhesion and colonization in the small intestine (Sheikh *et al.*, 2014). Another hypothetical protein encodes by gene *yaiT*, which has sequence similarity to the outer membrane porin family increases by 10-fold (Zhai & Saier, 2002).

The loss of *rpoS* does not affect expression of *yad* operon, while the expression is increased in double mutant of *rpoS hns* (Larsonneur *et al.*, 2016). The *yad* operon expression is depends on different environmental conditions including temperature, oxygen tension, pH, osmolarity, stringent response. Furthermore, the regulatory proteins of *E. coli*, including ArcA, Fnr, the two-component Cpx system, GadX and RpoS act as a repressor of *yad* operon. (Larsonneur *et al.*, 2016). In this current study, at 24h, *yad* operon was not expressed in WT, but expressed in *rpoS* mutants. However, at 48h, *yad* operon transcripts were highly abundant in WT and lower in *rpoS* mutants (Figure 21). This suggests the *yad* operon is negatively regulated by RpoS at 24h and positively at 48h of incubation. The data indicates that during prolonged-incubation RpoS controls growth phase-dependent gene expression (at least for putative fimbriae and adhesion genes).

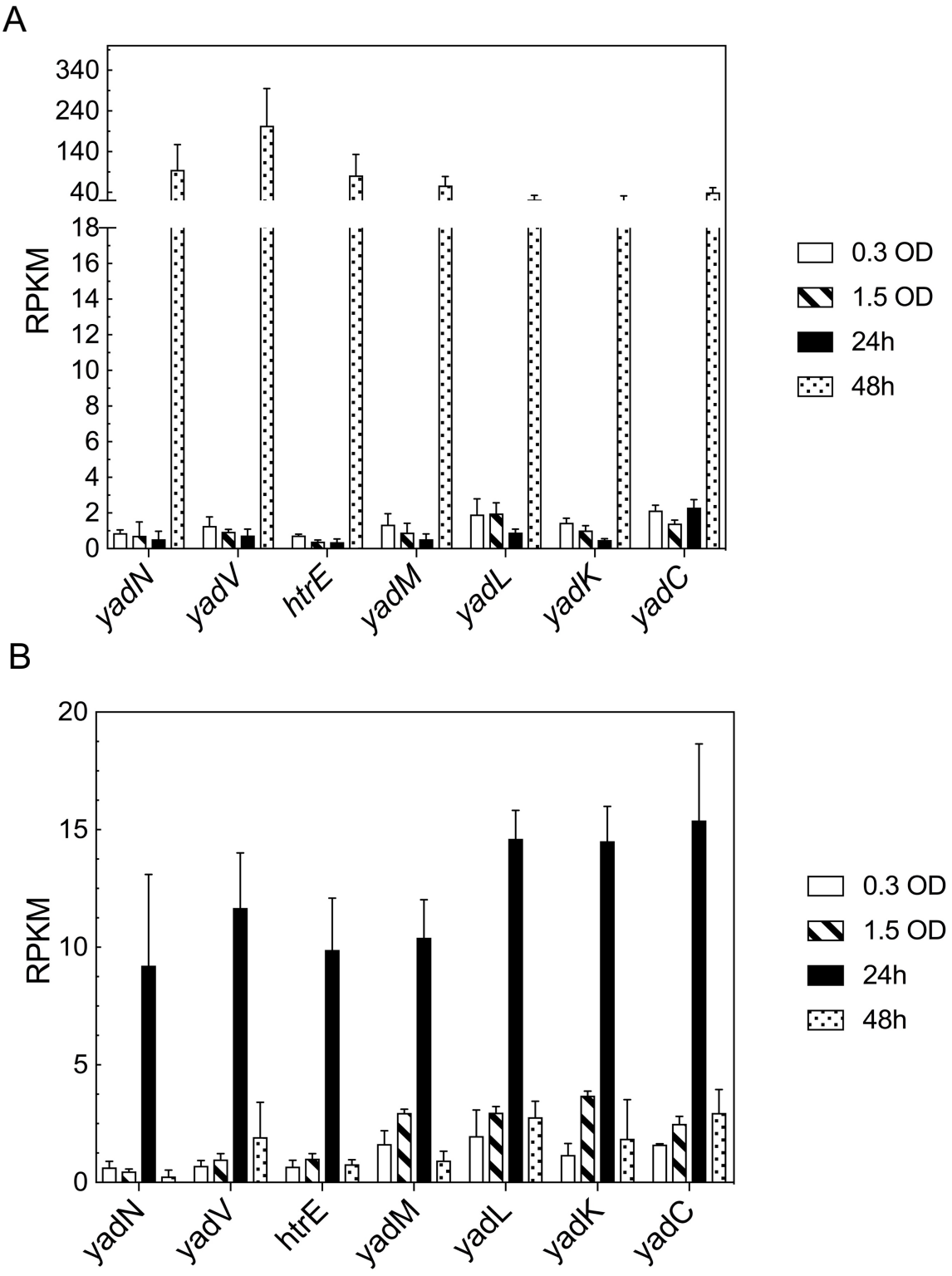


Figure 21: Putative fimbriae *yad* operon was highly expressed during prolonged-incubation in WT, while its expression was lower in Δ*rpoS* strain.
Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

Table 6: Transcript abundance of putative adhesion and fimbriae genes during prolonged-incubation in WT and Δ*rpoS* (48h relative to 24h of incubation).

			Fold-change (48h/24h)*	
Transcripts id	Gene	Protein/function	WT	Δ*rpoS*
AAC73252-1	*yadN*	Putative fimbrial protein	55.31	-16.22
AAC73251-1	*yadV*	Probable fimbrial chaperone	67.31	-5.55
AAC73250-1	*htrE*	Outer membrane usher protein	61.78	-10.55
AAC73249-1	*yadM*	Putative fimbrial-like protein	46.21	-8.75
AAC73248-1	*yadL*	Putative fimbrial-like protein	22.95	-5.46
AAC73247-1	*yadK*	Putative fimbrial-like protein	29.87	-6.99
AAC73246-1	*yadC*	Putative fimbrial-like protein	18.69	-5.39
AAC73632-1	*sfmA*	Putative fimbrial-like protein (Type-1A pilin)	12.90	-7.72
AAC73633-1	*sfmC*	Probable fimbrial chaperone	62.60	-6.50
AAC73634-1	*sfmD*	Outer membrane usher protein	28.95	-4.20
AAC73636-1	*sfmF*	Putative fimbrial-like protein	33.78	**-2.75**
AAC73635-1	*sfmH*	Putative fimbrial-like protein	61.79	-7.64
AAC75169-1	*yehA*	Putative fimbrial-like protein	13.28	-4.55
AAC75170-1	*yehB*	Outer membrane usher protein	33.57	-7.70
AAC75171-1	*yehC*	Probable fimbrial chaperone	34.28	-7.51
AAC73813-1	*ybgD*	Putative fimbrial-like protein	24.92	-5.78
AAC76177-1	*yraI*	Probable fimbrial chaperone	24.67	**-2.75**
AAC75393-1	*yfcP*	Putative fimbrial-like protein	7.79	**-2.45**
AAC75394-1	*yfcQ*	Putative fimbrial-like protein	5.79	**-1.47**
AAC75395-1	*yfcR*	Putative fimbrial-like protein	**2.85**	-5.14
AAC75396-1	*yfcS*	Probable fimbrial chaperone	8.05	-4.06
AAC75399-1	*yfcV*	Putative fimbrial-like protein	60.71	-4.59
AAC74026-1	*ycbR*	Probable outer membrane usher protein	7.28	**-2.87**
AAC74025-1	*ycbS*	Probable fimbrial chaperone protein	13.23	**-2.75**
AAC74028-1	*ycbU*	Putative fimbrial-like protein	12.53	**-2.23**
AAC74029-1	*ycbV*	Putative fimbrial-like protein	8.97	**-2.96**
AAC74575-1	*ydeQ*	Putative fimbrial-like protein	70.57	-14.58
AAC76247-1	*yhcA*	Putative fimbrial chaperone	54.60	-4.14
AAC76079-1	*ygiL*	Putative fimbrial-like protein	5.72	**-3.16**
AAC73395-1	*yagY*	Probable fimbrial chaperone	7.49	**-2.78**

Asterisk (*) indicates significant fold-change (FDR adjusted p-value ≤ 0.05). Significant fold-change less than 4-fold denoted as bold numbers. (-) sign indicates down-regulation of genes.

3.6 Toxin-Antitoxin transcripts

Toxin-Antitoxin systems (TA system) coding genes were highly expressed in WT during prolonged-incubation phase (Figure 22). Among these genes *mqsRA, yafO-yafN, dinJ-yafQ, yoeB-yefM* and *higBA* showed RpoS-independent expression. The role of toxin-antitoxins genes in biofilms and persistence has emerged recently. In terms of the genetic basis of persister formation, the main model holds that TA pairs are primarily responsible, as they induce dormancy (Jayaraman, 2008). MqsRA is the first TA system directly associated with persister cells formation in *E. coli*. Deletion of the *mqsRA* locus as well as *mqsR* alone, decreased persister formation, while production of MqsRA increased persistence (Kim & Wood, 2010). MqsR relies on Hha and CspD to form persister cells. Hha is another toxin paired with the antitoxin TomB that increases persister cell formation (Christensen-Dalsgaard *et al.*, 2010b). CspD is a stress-induced cold shock protein that is a DNA replication inhibitor as well as the stationary phase-specific gene that is induced independently of RpoS (Inouye, 1997). Additionally, *mqsR* is the most induced gene in persister cells (Shah *et al.*, 2006). Other toxin genes that are also highly induced in persister cells are *relE, higB, mazF, yafQ* and *yoeB* (Keren *et al.*, 2004). MqsR also act as a motility quorum sensing regulator gene that regulates motility-related promoters of genes during biofilm formation (Gonzalez Barrios *et al.*, 2006). Moreover, the anti-toxin MqsA directly represses the transcription of RpoS and reduces the concentration of c-di-GMP by repressing the diguanylate cyclases that are controlled by RpoS (Landini, 2009). In addition, *csgD*, which encodes the regulator for curli and cellulose, is activated by RpoS (Pesavento *et al.*, 2008), and repressed by MqsA. The result of repressing these RpoS regulated genes by MqsA leads to increased motility and a reduction in cell adhesion (Wang & Wood, 2011). Moreover, the toxin YafQ represses both RpoS and TnaA resulting in the reduction of indole levels, which leads to increased persistence. Levels of both

RpoS and TnaA are reduced when YafQ is overexpressed from a plasmid. In addition, indole levels are also lower when YafQ is overexpressed and persistence is higher. The role of indole in the repression of persistence is confirmed using a persistence assay (Kim *et al.*, 2010). The five TA systems (MazFE, RelEB, YoeB/YefM, YafQ/DinJ and ChpSB) are important for biofilm as deletion of each system decreased biofilm formation. The defect in biofilm formation is mainly a result of decreased cell lysis due to deletions in the toxin genes *mazF* and *yafQ (Kolodkin-Gal et al., 2009).* The programmed cell death could be an altruistic mechanism to allow a small subpopulation of biofilm cells to survive by releasing essential nutrients in the biofilm community where diffusion is limited. Interestingly*, mazEF* mediates cell death both in liquid media and during biofilm formation, while YafQ-DinJ is unique in that it is responsible for death process only during biofilm formation (Kolodkin-Gal *et al.,* 2009). The current data indicates high expression of the TA systems during prolonged-incubation, however, the role of TA systems during this phase remains to be tested. Furthermore, it will be interesting to study the signal that triggers the expression of the TA systems during prolonged-incubation.

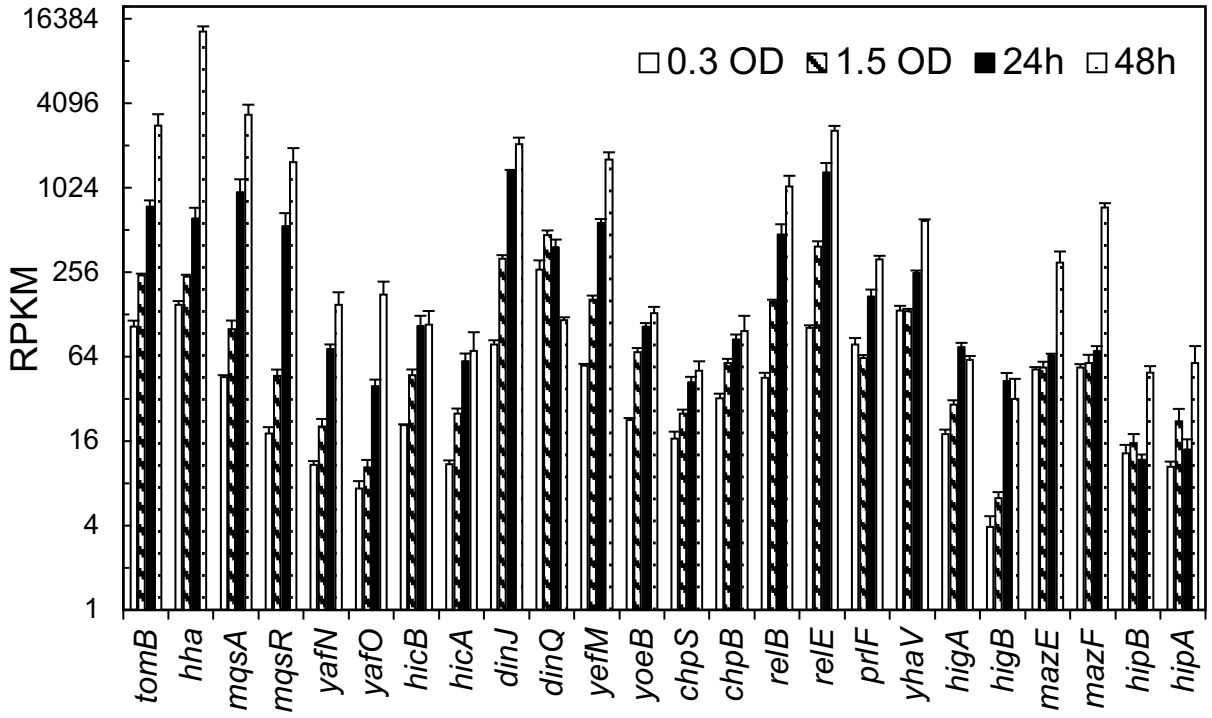


Figure 22: Toxin-antitoxin transcripts were in higher abundance during prolonged-incubation in WT.
Error bars represent standard error. RPKM indicates Reads Per Kilobase of Million mapped reads.

3.7 Biofilm-related transcripts

Many biofilm-related genes were expressed during prolonged incubation. At the initial phase of prolonged-incubation (at 24h relative to early stationary phase), specific genes which are tailored toward suppression of biofilm formation were preferentially expressed (Table 7) in WT. As discussed earlier, iron acquisition transcripts were high in abundance during this phase due to iron limitation and/or oxidation. In *E. coli*, iron regulation plays a critical role in biofilm formation. The addition of an iron chelator to cells growing in LB limits biofilm formation. This occurs through perturbation of 2Fe-2S homeostasis and subsequent activation of IscR. IscR is shown to control phase variation and fimbriae expression by inducing site-specific recombinase *fimE* which turns off fimbriae expression (Wu & Outten, 2009). Additionally, *fimE* gene was highly up-

regulated during initial phase of prolonged-incubation (Table 7). The major fimbriae structural gene *fimA* lies within an invertible DNA segment known as *fimS*. The orientation of the switch determines the transcription of the *fimA* gene. The two regulatory and site-specific recombinase proteins, FimB and FimE, catalyze the inversion process. Usually, expression of *fimE* gene orients the switch to the OFF orientation, while *fimB* gene expression turns the switch ON (Beloin *et al.*, 2008). During iron starvation, the iron-free Fur (*apo*-fur) is bound on the promoter region of the *ycgZ-ymgA-ariR-ymgC* operon and regulates the expression of this operon (Seo *et al.*, 2014). Three genes (*ymgA, ariR and ymgC*) in this operon are associated with biofilm formation and one of them (*ariR*) is also associated with acid resistance. The acid resistance regulator, *ariR* represses biofilm formation by repressing motility and provides acid resistance by acting as a non-specific DNA-binding protein (Lee *et al.*, 2007). Moreover, the *apo*-fur activation of *ymgA, ymgC* and *ariR* suppresses biofilm formation and enables planktonic growth of the cells, which allows them to find an iron-rich environment. Thus, Fur could play a key role in the suppression of biofilm formation and resistance to acidic stress by activating this particular operon during iron-limitation. In the current study, the transcripts level of *ycgZ-ymgA-ariR-ymgC* operon was strongly up-regulated and expressed as RpoS-independent transcripts (Table 7).

The *cpxP* (induced 2-fold at 24h relative to early stationary phase) gene encodes the periplasmic protein that mediates the response to envelop stress and many other cellular processes. *cpxR*-P additionally represses motility by downregulating flagellar gene expression. The *cpx* pathway also represses fimbriae expression which is necessary for the initial establishment of adhesion during biofilm formation. One of the inducing signals for the activation of the *cpx* pathway is alkaline pH. Since pH of LB after 24h is consistently found to be near 8 (as a result of degradation of amino acids which produce ammonium), elevated pH may be an inducing signal

for the *cpx* pathway (Dorel *et al.*, 2006). Expression of *dicF*, encoding a small regulatory RNA which represses RpoS expression and reduces biofilm formation and motility was increased by 5-fold at 24h relative to early stationary phase (Bak *et al.*, 2015). 2-fold enhanced expression was observed for the *ychH gene*, which encodes for stress-induced proteins and is activated in response to oxidative stress and serves to reduce biofilm formation (Lee *et al.*, 2010).

The gene *yliH/bssR* encode a global regulator protein that represses biofilm formation during growth in LB medium supplemented with glucose by increasing indole import, increasing catabolite repression by glucose import, and by decreasing AI-2 uptake (Domka *et al.*, 2006). In WT, the *bssR* gene was highly abundant during the prolonged-incubation phase (24h relative to early stationary phase). The high abundance of *bssR* transcript in overnight cultures is consistent with other study (Domka *et al.*, 2006). Interestingly, the *bssR* gene is negatively regulated by RpoS in rich media during the transition to stationary phase (Patten *et al.*, 2004). However, in *rpoS* mutants, the *bssR* transcripts were down-regulated during 24h of incubation relative to early stationary phase. This suggests growth phase-specific gene expression that agrees with the hypothesis that RpoS controls a distinct set of genes during prolonged-incubation, which are different from the RpoS-regulon members that are expressed during stationary phase.

Table 7: Biofilm-related transcripts for biofilm suppression were in high abundance during initial prolonged-incubation (24h relative to early stationary phase).

			Fold-change (24h/1.5)*	
Transcripts id	Gene	Protein/Function	WT	Δ*rpoS*
AAC77269-1	*fimE*	Type 1 fimbriae regulatory protein	4.86	11.45
AAC74248-1	*ycgZ*	Probable two-component-system connector protein	25.54	2.04
AAC74249-1	*ymgA*	Probable two-component-system connector protein	29.80	2.34
AAC74251-1	*ymgC*	Uncharacterized protein	24.77	6.13

Transcripts id	Gene	Protein/Function	Fold-change (24h/1.5)*	
			WT	Δ*rpoS*
AAC74250-1	*ariR*	Probable two-component-system connector protein	32.46	3.18
AAT48235-1	*cpxP*	Periplasmic protein	**2.59**	NS
EBE00001515049	*dicF*	Small regulatory protein	5.71	6.05
AAC74289-1	*ychH*	Uncharacterized protein	**2.56**	2.71
AAC73923-1	*bssR*	Biofilm regulator	30.76	-23.34

Asterisk (*) indicates significant fold-change (FDR adjusted p-value ≤ 0.05). If significant fold-change is less than 4-fold denoted as bold numbers. (-) sign indicates down-regulation of genes. NS for Non-significant.

Interestingly, at the later phase of prolonged-incubation (48h relative to 24h of incubation) the genes responsible for biofilm formation were up-regulated in WT (Table 8). During the early phase of biofilm development, adhesive organelles and curli fimbriae play a major role in attachment of cells to surfaces. Consistently, the genes for adhesion and putative fimbriae were strongly increase in expression during later phase of prolonged-incubation as discussed in RpoS-dependent transcripts. In addition, the expression level of *fimB* was high at 48h relative to 24h of incubation, which controls fimbriae ON and OFF switch. Expression of the FimB protein induces adhesion and initial development of biofilm (Beloin *et al.*, 2008). Furthermore, genes encoding minor structural proteins *fimFGH* were expressed along with outer membrane usher protein-coding gene *fimD*. NanR is a regulator that controls catabolism of *N*-acetyl-neuraminic acid (commonly sialic acid) and also controls fimbriae ON and OFF switch (Sohanpal *et al.*, 2004). NanR transcriptional repressor has 11 regulatees, 10 of which were induced at 48h compared to 24h of incubation (Figure 23), suggesting that NanR mediates the activation of FimB. However, it is not clear what signals initiate this program.

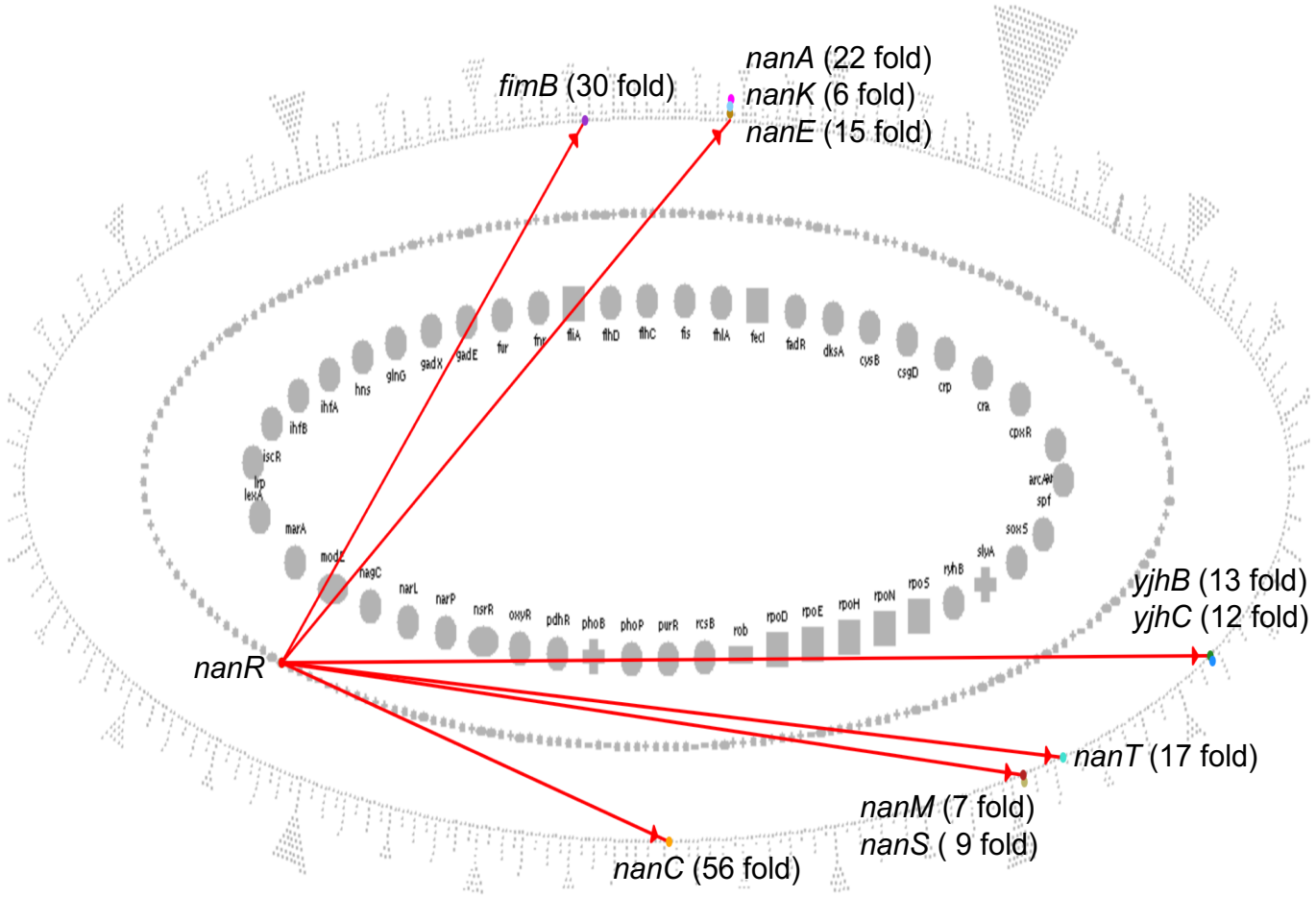


Figure 23: NanR and its regulatees were up-regulated during the later stage of prolonged-incubation (48h relative to 24h of incubation).
The regulatory overview of *E. coli* is presented here, where the outer ring page only those genes that are regulated, middle ring represents all other genes and inner ring is for master regulators and sigma factors (+/- are genes that have regulators/inhibitors only, round shapes are for genes that have activators and inhibitors, oval shapes are for genes that have all regulators but with unknown mode of regulation and square shapes are for sigma factors) (Generated by EcoCyc)

The genes for curli fimbriae synthesis were also expressed during the later phase of prolonged-incubation and also as RpoS-dependent. This includes *csgBAC* operon, which encodes the minor and major subunit of curlin protein and its assembly protein. The regulator of curli synthesis genes, *csgD* is the master transcriptional activator of CsgA protein and curli synthesis genes and plays a vital role in cell-to-surface attachment. Surprisingly, the transcripts levels of *csgD* were elevated during the initial phase of prolonged-incubation. This suggests that regulation of curli fimbriae genes is initiated at the start of prolonged-incubation and later controls the expression of the curli fimbriae genes. Furthermore, the operon *pgaABCD* is required for synthesis, modification and export of cell-bound hexosamine-rich polysaccharide, known as β-1,6-N-acetylglucosamine (PGA), adhesins essential for biofilm formation (Itoh *et al.*, 2008). The genes *pgaD and pgaC* codes for glycosyltransferase which facilitates the export and localization of PGA polymers to the periplasm (Vogeleer *et al.*, 2014). In *E. coli* K-12 strains the expression of these genes are significant for cell-to-cell adhesion and attachment to surfaces (Agladze *et al.*, 2005). The operon was highly expressed during the later phase of prolonged-incubation and expressed as RpoS-dependent (Table 8).

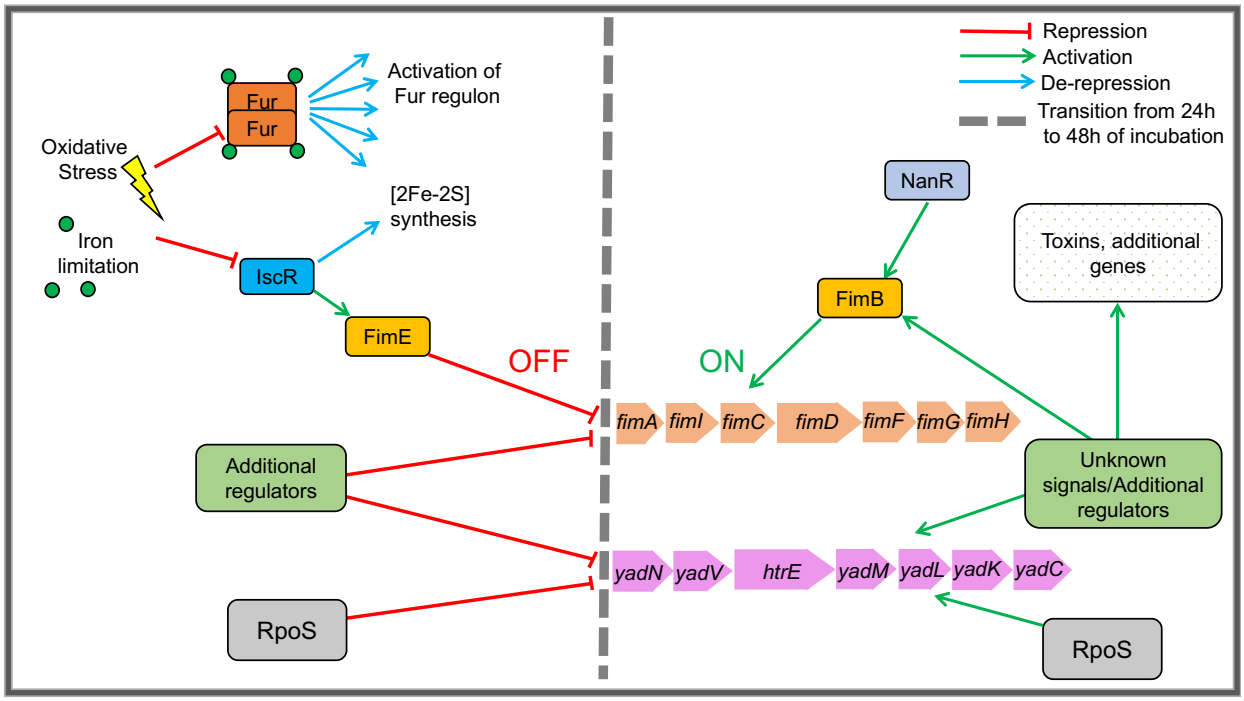


Figure 24: Predicted regulation of fimbria and putative fimbriae genes during prolonged-incubation.
FimE is induced during 24h of incubation that repressed the *fim* operon. However, during 48h of incubation FimB is highly expressed, which in turn activates the *fim* operon. NanR may be an inducing signal, which activates the expression of *fimB* gene. RpoS serves to limit the expression of *yad* operon at 24h of incubation, while at 48h of incubation RpoS might require for the expression of the *yad* operon.

An environmentally responsive signal transduction system that controls expression and/or activity of the enzymes GCDEF and EAL domain-containing proteins are responsible for synthesis and degradation of c-di-GMP (Povolotsky & Hengge, 2012). The genes for diguanylate cyclases (GGC) enzymes encoded by *ycdT* (*dgcT), yeaJ(dgcJ), and yliF (dgcI)* were in strongly expressed during the later phase of prolonged-incubation (Table 9). c-di-GMP, an allosteric activator is observed to promote biofilm formation and synthesis of adhesins (Cotter & Stibitz, 2007). The regulatory proteins, Fis and CRP activate the expression of the *yeaJ* gene (Amores *et al.*, 2017). *dgcT* gene is also implicated in the production of poly-GlcNAc, which serves as a biofilm matrix component and/or virulence factor in some pathogenic *E. coli* (Jonas *et al.*, 2008). As mention in

the discussion of Toxin-antitoxin systems, the TA system also plays a positive role in biofilms formation (Kolodkin-Gal *et al.*, 2009, Ren *et al.*, 2004, Gonzalez Barrios *et al.*, 2006, Harrison *et al.*, 2009), which further supports the hypothesis that the biofilm responsible genes are expressed during the later phase of prolonged-incubation. Altogether, the data suggest that in planktonic cultures during prolonged-incubation, biofilm-related genes are expressed. Initially, genes expressed are tailored towards suppression, while at later in the prolonged-incubation phase genes that promote biofilm formation are expressed. The role of RpoS in the regulation of these genes is conflicting as both RpoS-independent and dependent genes expression are observed.

Table 8: Abundance of transcripts responsible for biofilm formation during the later phase of prolonged-incubation.

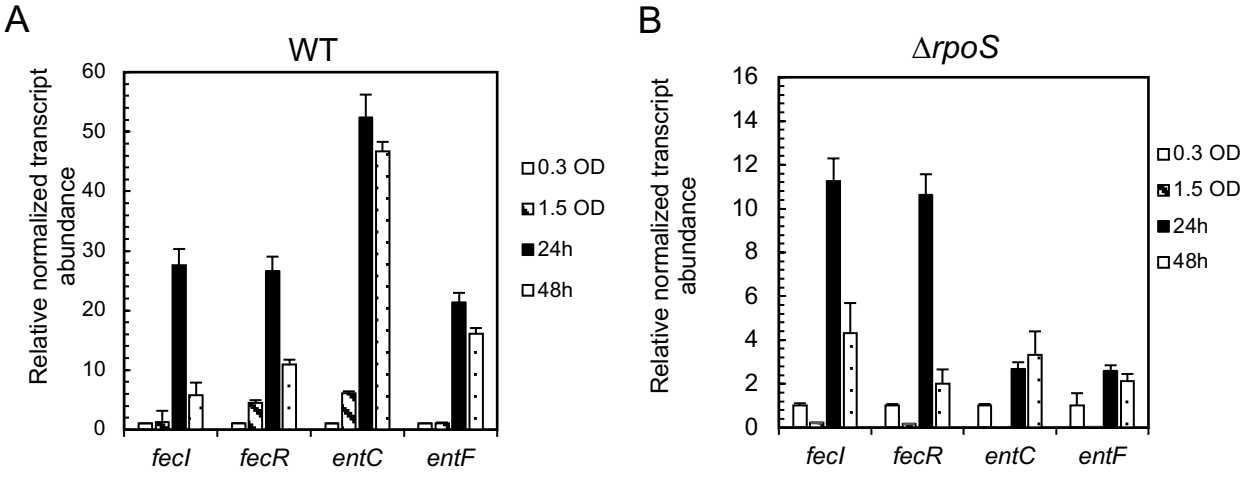
			Fold-change (48h/24h)*	
Transcripts id	Gene	Protein/Function	WT	Δ*rpoS*
AAC77268-1	*fimB*	Type 1 fimbriae regulatory protein	29.54	NS
AAC77273-1	*fimD*	Outer membrane usher protein	13.81	NS
AAC77274-1	*fimF*	Type-1 fimbrial minor subunit	5.17	NS
AAC77275-1	*fimG*	Type-1 fimbrial minor subunit	**3.17**	NS
AAC77276-1	*fimH*	Type 1 fimbrin D-mannose specific adhesin	4.22	NS
AAC76257-1	*nanA*	N-acetylneuraminate lyase	22.65	5.80
AAC77267-1	*nanC*	Probable N-acetylneuraminic acid outer membrane channel protein	55.93	-8.37
AAC76255-1	*nanE*	Putative N-acetylmannosamine-6-phosphate 2-epimerase	15.39	**2.92**
AAC76254-1	*nanK*	N-acetylmannosamine kinase	5.93	**3.51**
AAC77266-1	*nanM*	N-acetylneuraminate epimerase	6.91	**-3.20**
AAC77265-1	*nanS*	Probable 9-O-acetyl-N-acetylneuraminic acid deacetylase	9.36	NS
AAC76256-1	*nanT*	Sialic acid transporter	17.83	4.23
AAC77235-1	*yjhB*	Putative metabolite transport protein	13.29	**3.94**
AAC77236-1	*yjhC*	Putative oxidoreductase	12.73	**2.95**
AAC74126-1	*csgA*	Major curlin subunit	**2.95**	NS
AAC74125-1	*csgB*	Minor curlin subunit	15.79	-12.24
AAC74127-1	*csgC*	Curli assembly protein	6.97	**-3.98**
AAC74109-1	*pgaA*	Poly-beta-1,6-N-acetyl-D-glucosamine export protein	13.47	NS

Transcripts id	Gene	Protein/Function	Fold-change (48h/24h)*	
			WT	Δ*rpoS*
AAC74108-1	*pgaB*	Poly-beta-1,6-N-acetyl-D-glucosamine deacetylase	28.35	**-3.26**
AAC74107-1	*pgaC*	Poly-beta-1,6-N-acetyl-D-glucosamine synthase	19.24	-5.17
AAC74106-1	*pgaD*	Biofilm PGA synthesis protein	9.23	NS
AAC74110-1	*ycdT*	Probable diguanylate cyclase (DGC)	41.73	NS
AAC74856-1	*yeaJ*	Putative diguanylate cyclase (DGC)	17.26	NS

Asterisk (*) indicates significant fold-changes ≥ 4 and FRD adjusted *p* value ≤ 0.05. If fold changes are less than 4-fold are denoted as bold numbers. (-) sign indicates down-regulation of genes. NS for Non-significant.

3.8 Validation of gene expression data using qPCR

To validate the gene expression data, the expression of a set of representative genes was further examined by RT-qPCR. Four genes from the iron acquisition group *fecI, fecR, entC* and *entF*, two from degradation enzyme-coding genes *astA* and *astC*, and two from Toxin-antitoxin system *mqsR* and *mqsA* genes were selected. Their RpoS-independent expression during prolonged-incubation was confirmed using RT-qPCR (Figure 25). Moreover, three genes coding putative fimbrial protein *yadN, yadV* and *sfmH* were selected and their RpoS-dependent expression was confirmed (Figure 26).



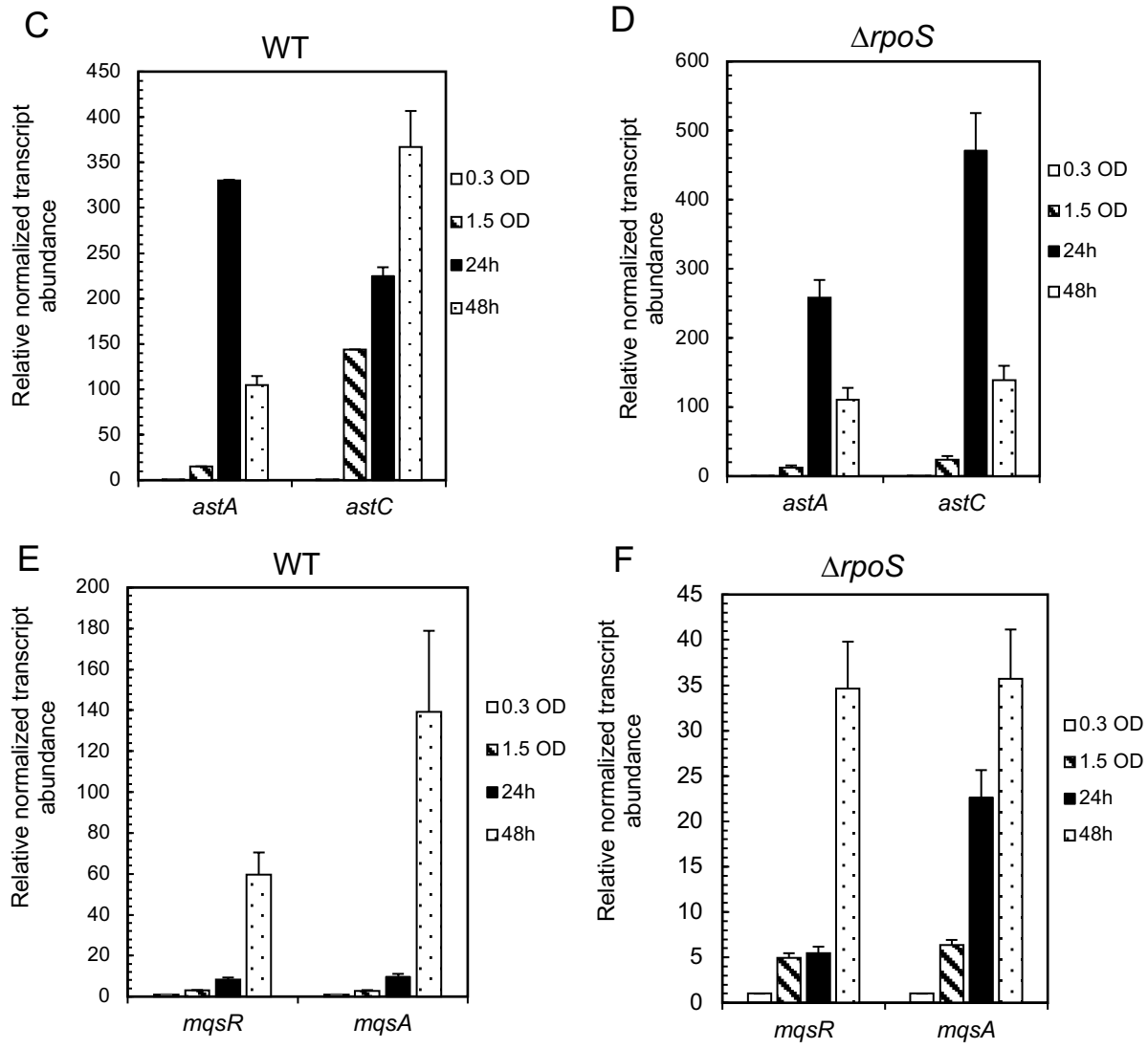


Figure 25: RpoS-independent gene expression during prolonged-incubation were validated using RT-qPCR.

The transcript abundance of iron acquisition genes in WT (A) and Δ*rpoS* (B), transcript abundance of arginine degradation protein-coding genes in WT (C) and Δ*rpoS* (D) and transcript abundance of toxin-antitoxin protein-conding genes in WT (E) and Δ*rpoS* (F). The transcript abundance was normalized using 16s rRNA gene (*rrsA*) and relative to exponential phase (0.3 OD). Error bars represent standard error.

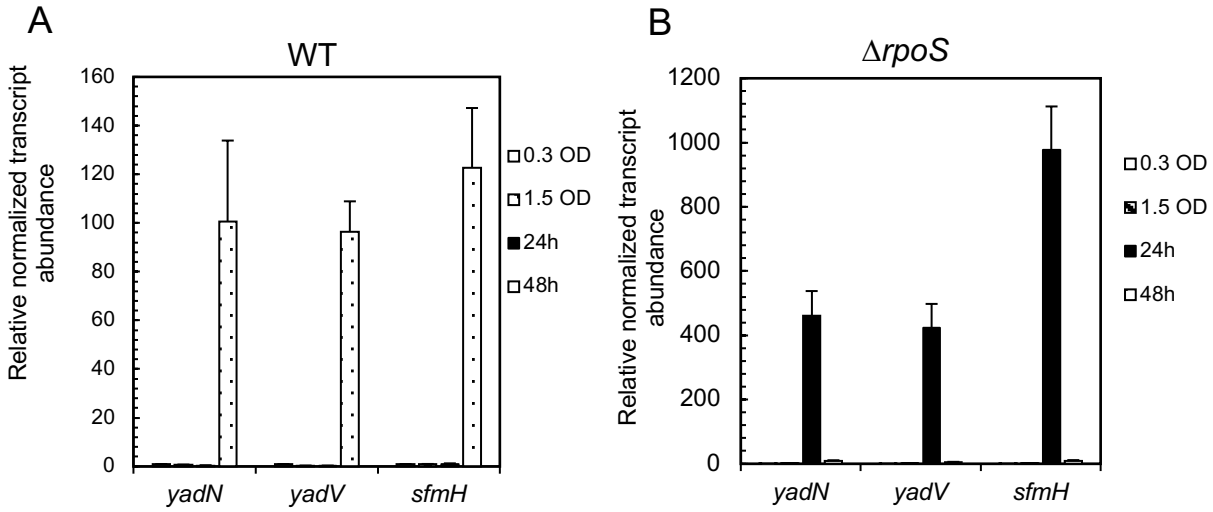


Figure 26: RpoS-dependent gene expression during prolonged-incubation were validated using RT-qPCR.
The transcript abundance of putative fimbriae and adhesion protein-coding genes in WT (A) and Δ*rpoS* (B). The transcript abundance was normalized using 16s rRNA gene (*rrsA*) and relative to exponential phase (0.3 OD). Error bars represent standard error.

In summary, the current study indicates that during prolonged-incubation, a unique set of both RpoS-independent and dependent genes are expressed, which agrees with the hypothesis that along with RpoS, other regulators may play an essential role in controlling gene expression during prolonged-incubation. The RpoS-independent genes are not observed as simple linear increment as predicted, may be due to dynamic changes in transcript abundance in cells and even in growth arrest cells. RpoS control genes during prolonged-incubation that are distinct from the stationary phase expressed genes. Surprisingly, genes required for biofilm formation are also expressed in planktonic cultures during prolonged-incubation. The genes identified in this study should help and guide future efforts to uncover how physiological adaptation mediated by the identified genes helps *E. coli* to survive prolonged starvation.

3.9 Future directions

The current genomic expression profiling study on prolonged-incubation phase has expanded our understanding on genes expressed during prolonged starvation. Nevertheless, we still do not know their importance in survival of *E. coli* cells. For that reason, a potential focus area is the physiological effects of prolonged-incubation phase genes in *E. coli* cells on their viability and survival. Tn-seq (Transposon mutagenesis pair with massively parallel sequencing) can be used to identify genes that are essential for survival. Also, mutants of select prolonged-incubation phase genes can be generated to test for viability and survival during prolonged starvation conditions. Furthermore, quantitative analysis of proteins using SILAC (Stable Isotope Labelling with Amino acids in Cell culture) will provide insight on dynamics of the proteome in *E. coli* cells during prolonged-incubation phase. The current study was done on the laboratory strain of *E. coli* (MG1655 K12), however relevance to pathogenicity and natural environment survival is not clear. This could be explored using well known pathogenic strains (For instance: EDL933 O157:H7) and defined collection of natural *E. coli* isolates, which will provide a comprehensive understanding of the survival mechanisms in hosts and in natural environments.

APPENDIX 1: RNA integrity and quality check

Table 9: Yield and quality of RNA isolated from WT (*E. coli* K12 MG1655) in exponential (OD600 = 0.3) and early stationary (OD600 = 1.5), prolonged incubation (24h and 48h) phases of growth.

Growth phase	Replicate no.	Total amount (ug)	Volume(µl)	Yield (ug)	RNA recovered (%)	Abs 260/280
0.3 OD	1	14.0	35.0	9.0	64.5	2.1
	2	16.2	35.0	11.2	69.0	2.1
	3	20.3	35.0	11.9	58.7	2.1
1.5 OD	1	13.2	35.0	7.6	58.0	2.0
	2	10.3	35.0	7.9	77.1	2.1
	3	16.8	35.0	9.4	55.9	2.0
24h	1	11.1	35.0	6.6	59.7	2.0
	2	12.8	35.0	8.0	62.7	2.1
	3	10.0	35.0	8.3	83.1	2.1
48h	1	7.0	35.0	4.1	58.1	2.0
	2	9.1	35.0	4.6	50.7	2.1
	3	8.5	35.0	4.8	56.6	2.0

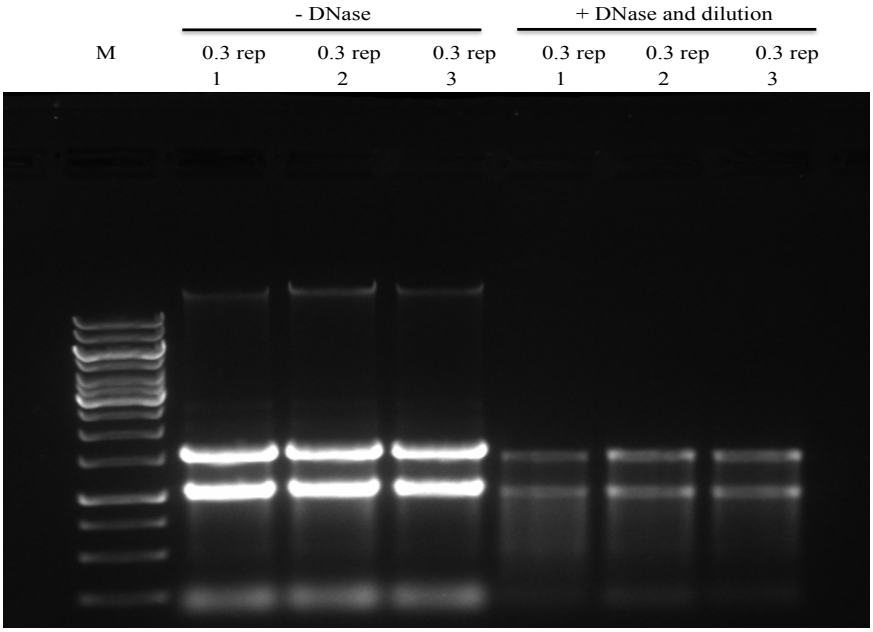


Figure 27: RNA integrity check for 0.3 OD samples in WT.
A 3-µl aliquot of RNA was run on a 0.8% agarose gel to check for integrity. The samples were approximately 5-fold diluted, as the RNA concentration were high for exponential phase samples.

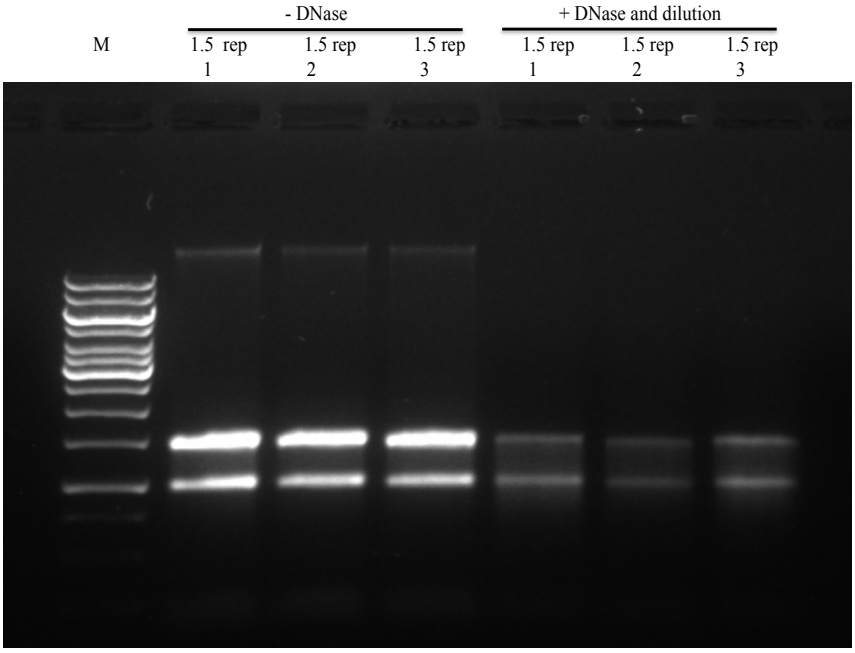


Figure 28: RNA integrity check for 1.5 OD samples in WT.
A 3-µl aliquot of RNA was run on a 0.8% agarose gel to check for integrity. The samples were approximately 2-fold diluted.

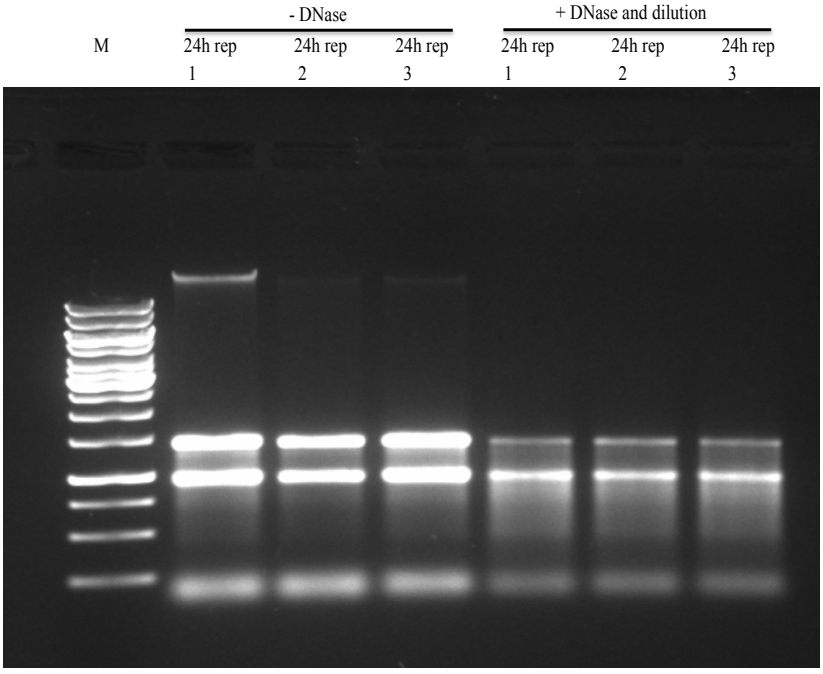


Figure 29: RNA integrity check for 24h samples in WT.
A 3-µl aliquot of RNA was run on a 0.8% agarose gel to check for integrity. The samples were approximately 2-fold diluted.

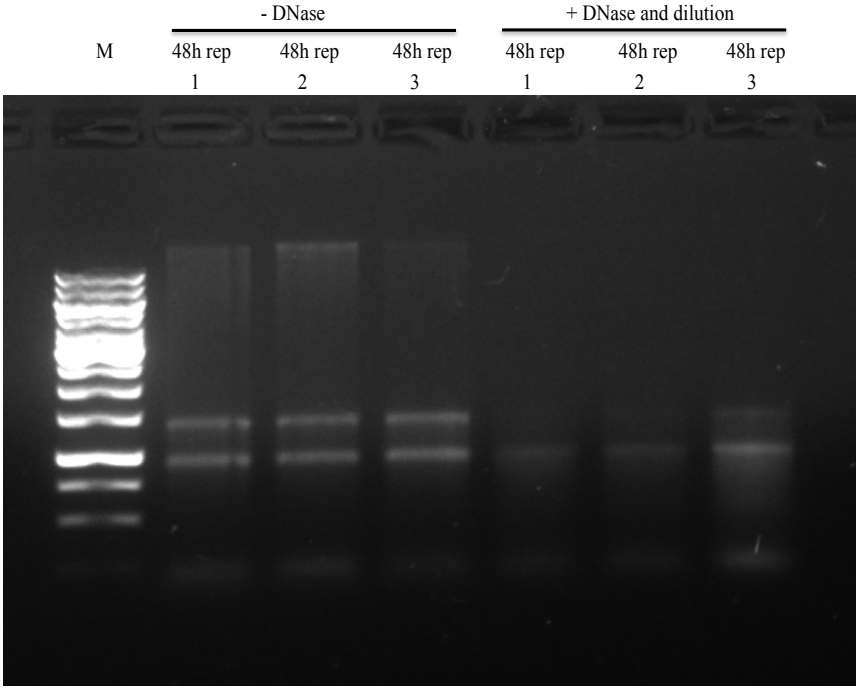


Figure 30: RNA integrity check for 48h samples in WT.
A 3-µl aliquot of RNA was run on a 0.8% agarose gel to check for integrity. The samples were approximately 1.25-fold diluted, as the RNA concentration were low for 48h samples.

Table 10: Yield and quality of RNA isolated from Δ*rpoS* (isogenic mutant of *E. coli* K12 MG1655) in exponential (OD600 = 0.3) and early stationary (OD600 = 1.5), prolonged incubation (24h and 48h) phases of growth.

Growth phase	Replicate no	Total amount (ug)	Volume(µl)	Yield(ug)	% RNA recovered	Abs 260/280
	1	11.0	35	7.4	67.3	2.1
0.3 OD	2	17.5	35	8.8	50.4	2.1
	3	16.0	35	7.3	45.6	2.1
	1	10.4	35	6.2	59.6	2.1
1.5 OD	2	11.6	35	7.1	60.9	2.1
	3	13.2	35	7.3	55.8	2.1
	1	15.6	35	9.6	61.6	2.1
24h	2	11.1	35	6.4	57.8	2.1
	3	10.1	35	7.5	74.5	2.1
	1	7.2	35	4.9	68.2	2.1
48h	2	8.7	35	6.2	71.9	2.1
	3	7.1	35	3.5	50.2	2.1

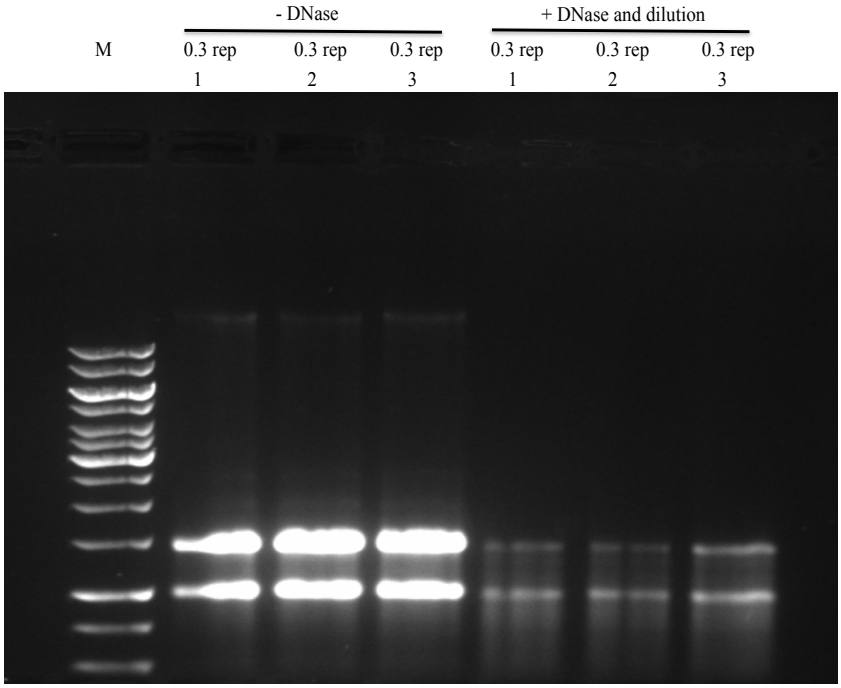


Figure 31: RNA integrity check for 0.3 OD samples in Δ*rpoS*.
A 3-µl aliquot of RNA was run on a 0.8% agarose gel to check for integrity. The samples were approximately 5-fold diluted, as the RNA concentration were high for exponential phase samples.

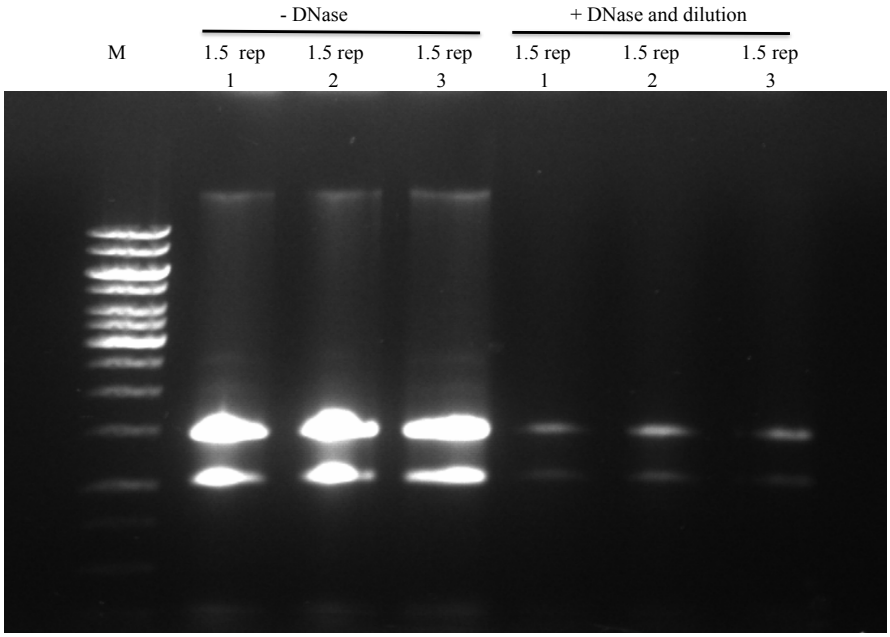


Figure 32: RNA integrity check for 1.5 OD samples in Δ*rpoS*.
A 3-µl aliquot of RNA was run on a 0.8% agarose gel to check for integrity. The samples were approximately 2-fold diluted.

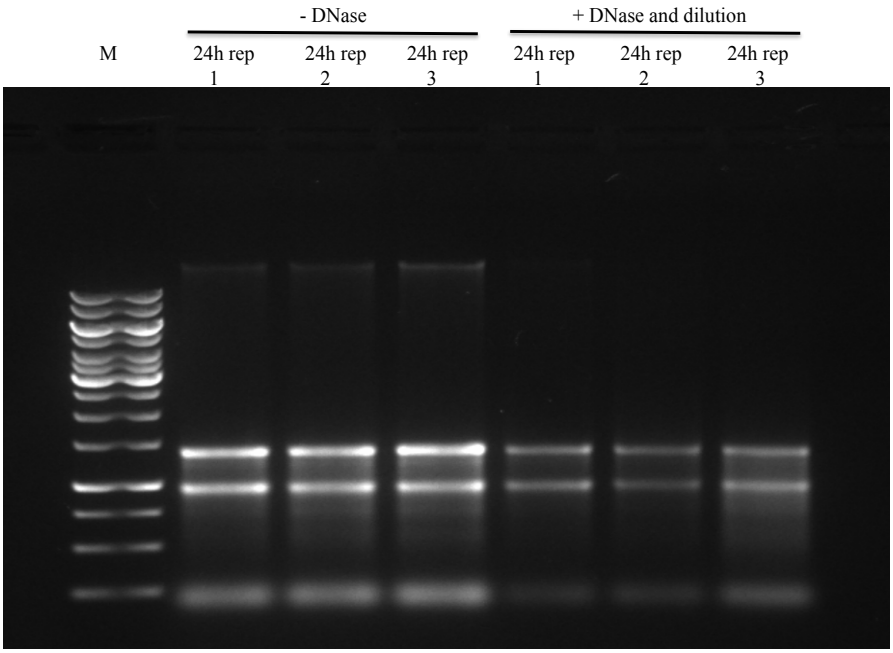


Figure 33: RNA integrity check for 24h samples in Δ*rpoS*.
A 3-µl aliquot of RNA was run on a 0.8% agarose gel to check for integrity. The samples were approximately 2-fold diluted.

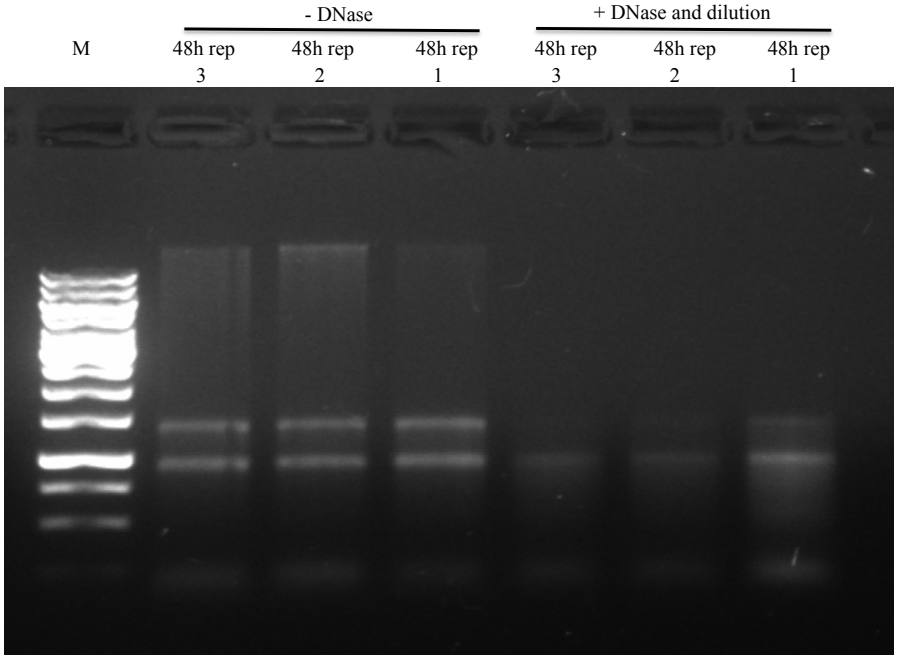


Figure 34: RNA integrity check for 48h samples Δ*rpoS*.
A 3-µl aliquot of RNA was run on a 0.8% agarose gel to check for integrity. The samples were approximately 1.25-fold diluted, as the RNA concentration were low for 48h samples.

APPENDIX 2: Total percentage of mapped reads in RNA sequencing data

Table 11: The total generated reads in RNA-sequencing and total mapped reads using Bowtie2 for WT and Δ*rpoS*.

Strain	Growth phase	Replicate number	Number of Reads HiSeq	Number of mapped reads	Percentage of alignment
WT	0.3 OD	1	21,315,490	21,126,143	99.1
		2	18,792,038	18,525,994	98.5
		3	21,653,685	21,475,077	99.2
	1.5 OD	1	24,141,417	23,724,898	98.3
		2	20,937,041	20,601,476	98.4
		3	20,179,295	19,909,837	98.7
	24h	1	19,233,834	18,424,594	95.8
		2	13,963,416	13,614,261	97.5
		3	18,036,070	17,430,380	96.6
	48h	1	22,046,420	3,345,543	15.2
		2	18,920,328	16,737,471	88.5
		3	23,289,075	3,340,655	14.3
rpoS mutant	0.3 OD	1	20,334,340	20,201,105	99.3
		2	23,137,894	22,949,183	99.2
		3	19,457,519	19,312,437	99.3
	1.5 OD	1	23,029,307	22,801,642	99.0
		2	23,010,010	22,636,208	98.4
		3	21,138,226	20,878,283	98.8
	24h	1	17,070,442	16,567,736	97.1
		2	17,524,844	16,709,794	95.4
		3	15,407,051	14,948,325	97.0
	48h	1	17,318,094	3,638,263	21.0
		2	20,079,002	2,933,942	14.6
		3	19,878,503	17,169,896	86.4

Exponential phase-0.3 OD, Early stationary phase-1.5 OD and Prolonged-incubation phase-24h and 48h

APPENDIX 3: Total number of annotated transcripts used in correlation analysis

Table 12: The total number of annotated transcripts generated by each technique as well as those transcripts that were annotated by both techniques.

Sample information			RNA-sequencing data				Microarray data	
Strain	Growth phase	Rep. No.	Number of generated reads	Total no of reads after Normalization (RPKM)	Total no of annotated transcripts in RNA-seq	Common annotated transcripts in both technique *	Total no of annotated transcripts in Microarray	Total signal intensities in each sample after GCRMA normalization
WT	0.3 OD	1	21,315,490	1,426,659.34	4164	4010	4137	26,320.48
		2	18,792,038	1,413,338.19				26,178.50
		3	21,653,685	1,454,868.95				26,292.97
	1.5 OD	1	24,141,417	1,602,256.65				26,614.18
		2	20,937,041	1,594,937.81				26,820.28
		3	20,179,295	1,608,544.33				26,707.39
	24h	1	19,233,834	2,048,108.17				25,927.94
		2	13,963,416	2,223,436.27				25,359.72
		3	18,036,070	2,102,565.21				25,722.89
	48h	1	22,046,420	2,590,624.24				26,320.48
		2	18,920,328	2,339,251.98				26,320.48
		3	23,289,075	2,540,428.14				24,449.79
Δ*rpoS*	0.3 OD	1	20,334,340	1,365,959.09	4164	4010	4137	26,154.99
		2	23,137,894	1,452,267.72				26,131.34
		3	19,457,519	1,394,481.69				26,272.40
	1.5 OD	1	23,029,307	1,565,637.43				26,178.10
		2	23,010,010	1,498,924.25				26,100.61
		3	21,138,226	1,488,360.32				26,325.61
	24h	1	17,070,442	2,472,042.90				23,523.44
		2	17,524,844	2,298,312.37				24,474.60
		3	15,407,051	2,471,621.34				25,254.59
	48h	1	17,318,094	1,993,973.93				24,881.46
		2	20,079,002	1,871,523.21				24,627.29
		3	19,878,503	1,782,107.06				26,392.48

Asterisk (*) indicates common annotated transcripts in both the techniques are same for all growth phases and in both strains. 0.3 OD is Exponential phase, 1.5 OD is Early stationary phase, 24h and 48h is Prolonged-incubation

APPENDIX 4: MA plot of differentially-expressed genes in RNA-seq data

The overall summary for differentially-expressed transcripts in RNA-seq can be visualized through MA plots (Figure 35). This plot represents each gene with a dot. The *x*-axis is the average expression over the mean of normalized counts (A values), the *y*-axis is the $\log_2$ fold change between the provided condition (M values). Red dots indicate significantly altered (FDR adjusted $p \leq 0.05$) gene expression between the specified condition. Grey dots indicate no significant change in expression. At 24h of incubation relative to early stationary phase, the up-regulated genes were fewer in WT, but in Δ*rpoS* the number of up-regulated genes increases (Figure 35: A and C and also mention in Table 5). Thus, it can be postulated that RpoS negatively regulates many genes at initial phase of prolonged-incubation. Moreover, the correlation plot of gene expression data for 24h of incubation in WT (Figure 5: 24h) shows many low abundance transcripts while in Δ*rpoS* (Figure 6: 24h) very few low abundance transcripts are present, which further supports the hypothesis. Interestingly, at 48h relative to 24h the opposite result is observed, as the up-regulated genes are greater in WT (Figure 35: B) compared to Δ*rpoS*. Altogether, the data indicates that RpoS may play a differential role during prolonged-incubation, as negative regulation is observed at least for initial phase of prolonged-incubation (24h relative to early stationary phase) and positive regulation at later phase of prolonged-incubation (48h relative to 24h of incubation).

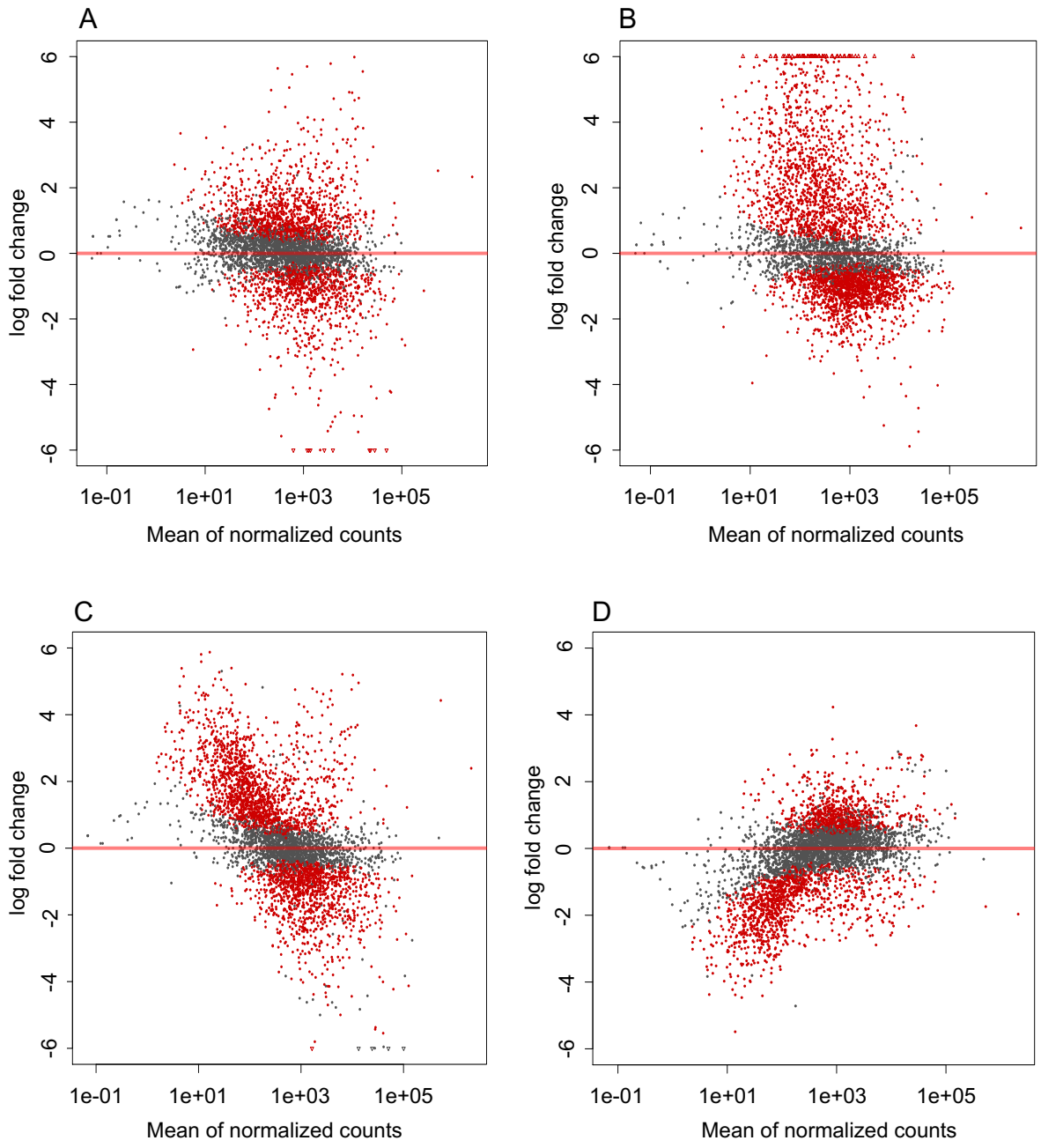


Figure 35: Global transcriptomic profile during prolonged-incubation in WT and ΔrpoS.
24h incubation compared to early stationary phase in WT (A) and in Δ*rpoS* (C) and 48h compared to 24h of incubation in WT (B) and in Δ*rpoS* (D). The y-axis represents the log2 fold-change in gene expression and the x-axis represents the mean read counts for each gene between the samples. The red dots represent the genes that are upregulated (above the red line) and downregulated (below the red line) with significantly altered expression values (FDR adjusted p-value ≤ 0.05).

APPENDIX 5: Complete list of transcripts that were higher in abundance during prolonged-incubation (24h relative to early stationary phase)

Table 13: List of transcripts that were higher in abundance during 24h relative to early stationary phase (OD600 =1.5).

Transcript id	Gene	Protein/Function	Fold-change (24h/1.5 OD) in WT		Fold-change (24h/1.5 OD) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC73686-1	*fes*	Enterochelin esterase	35.8	38.0	15.9	29.9
ABD18719-1	*yjjZ*	Uncharacterized protein	33.5		12.7	
AAC74250-1	*ariR*	Probable two-component-system connector protein	32.5	37.5		
ABD18641-1	*ybdZ*	Enterobactin biosynthesis protein	31.8		14.2	
AAC73923-1	*bssR*	Biofilm regulator	30.8	25.9	-23.34	-51.95
AAC74249-1	*ymgA*	Probable two-component-system connector protein	29.8	52.5		
AAC75584-1	*iscR*	HTH-type transcriptional regulator	26.5		4.4	4.4
AAC74248-1	*ycgZ*	Probable two-component-system connector protein	25.5			
AAC77020-1	*pspG*	Phage shock protein G	25.4		4.1	5.8
AAC76195-1	*mtr*	Tryptophan-specific transport protein	25.4			
AAC74251-1	*ymgC*	Uncharacterized protein	24.8		6.1	30.2
AAT48235-1	*cpxP*	Periplasmic protein	24.2			
AAC74817-1	*astA*	Arginine N-succinyltransferase	24.1	26.8	18.6	9.4
AAC74816-1	*astD*	N-succinylglutamate 5-semialdehyde dehydrogenase	24.0	29.3	13.6	18.3
AAC76710-1	*ibpA*	Small heat shock protein	24.0			
AAC76494-1	*zntA*	Zinc/cadmium/lead-transporting P-type ATPase	22.4		4.4	
AAC73687-1	*entF*	Enterobactin synthase component F	22.2		11.5	
AAC74818-1	*astC*	Succinylornithine transaminase	21.8	24.9	22.9	10.5
AAC76709-1	*ibpB*	Small heat shock protein	21.5		13.7	12.0
AAT48142-1	*iscS*	Cysteine desulfurase	21.3		4.4	4.1

Transcript id	Gene	Protein/Function	Fold-change (24h/1.5 OD) in WT		Fold-change (24h/1.5 OD) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC75582-1	*iscU*	Iron-sulfur cluster assembly scaffold protein	20.6		4.6	4.1
AAC75721-1	*nrdI*	Dimanganese-tyrosyl radical cofactor maintenance flavodoxin	19.9			
AAC74289-1	*ychH*	Uncharacterized protein	19.6			
AAC74815-1	*astB*	N-succinylarginine dihydrolase	19.3	46.7	9.8	4.0
AAC75720-1	*nrdH*	Glutaredoxin-like protein	18.6	23.4	8.4	
AAC75581-1	*iscA*	Iron-binding protein	17.9		4.2	
AAC77249-1	*fecI*	RNA polymerase sigma factor	17.2	35.0	4.2	
AAC75451-1	*mntH*	Divalent metal cation transporter	16.9	26.2	4.5	
AAC74814-1	*astE*	Succinylglutamate desuccinylase	16.6	33.6	10.1	
ABP93445-1	*yneM*	Uncharacterized protein	16.6		17.8	
AAC73696-1	*entB*	Enterobactin synthase component B	15.8	18.07	21.9	16.09
AAC77039-1	*acs*	Acetyl-coenzyme A synthetase	14.9	12.6	7.4	5.6
ACO60005-1	*yqeL*	Uncharacterized protein	14.9		12.0	
AAC74090-1	*ymdF*	Uncharacterized protein	14.6			
AAC74341-1	*yciG*	Uncharacterized protein	14.5	10.5	6.8	
AAC74683-1	*fumC*	Fumarate hydratase class II	14.5	11.8		
AAC77248-1	*fecR*	Regulator for fec operon	14.3	20.3	4.8	
AAC76767-1	*asnA*	Aspartate--ammonia ligase	14.3	24.3		
AAC76058-1	*mqsR*	mRNA interferase toxin (Motility quorum-sensing regulator)	14.2	15.4	4.3	13.5
AAC73695-1	*entE*	Enterobactin synthase component E	13.8		21.3	
ABD18643-1	*kdpF*	Potassium-transporting ATPase KdpF subunit	13.4		13.1	
AAC77199-1	*mgtA*	Magnesium-transporting ATPase	13.1		27.2	
AAC73694-1	*entC*	Isochorismate synthase	12.6		19.4	18.2
b2850-2	*ygeF*	Protein YgeF	12.4		19.9	40.2

Transcript id	Gene	Protein/Function	Fold-change (24h/1.5 OD) in WT		Fold-change (24h/1.5 OD) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC76057-1	*mqsA*	Antitoxin	12.1	12.8	4.6	7.0
AAC75706-1	*csiD*	PF08943 family protein	12.1	9.3		
AHA50631-1	*mntS*	Small protein	12.1			
AAC73434-1	*prpB*	2-methylisocitrate lyase	11.8	23.1	11.1	
AAC73558-1	*ybaA*	Uncharacterized protein	11.7	4.4		
AAC77322-1	*bglJ*	Transcriptional activator protein	11.6	26.9	8.8	
AAC73267-1	*erpA*	Iron-sulfur cluster insertion protein	11.6			
ABD18704-1	*yibT*	Uncharacterized protein	11.4			
AAC77323-1	*fhuF*	Ferric iron reductase protein	11.1		9.4	21.5
AAC74561-1	*ddpX*	D-alanyl-D-alanine dipeptidase	11.0	5.3	11.3	6.9
AAC73436-1	*prpC*	2-methylcitrate synthase	11.0		10.6	
AAC75722-1	*nrdE*	Ribonucleoside-diphosphate reductase 2 subunit alpha	11.0			
AAC74689-1	*uidA*	Beta-glucuronidase	11.0	10.5		
AAC76890-1	*sodA*	Superoxide dismutase	10.7	9.7	24.5	14.7
AAC74899	*htpX*	Protease (Heat shock protein)	10.7			
AAC73437-1	*prpD*	2-methylcitrate dehydratase	10.5		9.0	
AAC73608-1	*allR*	HTH-type transcriptional repressor	10.5			
EBE00001514986	*ileX*	tRNA	10.3		17.4	
AAC74244-1	*iraM*	Anti-adapter protein	10.3		13.0	50.1
AAC73697-1	*entA*	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase	10.1	9.9	11.3	
AAC74619-1	*tfaQ*	Prophage tail fiber assembly protein homolog	10.0	18.6		
AAC74688-1	*uidB*	Glucuronide carrier protein homolog	10.0			
AAC74456-1	*pinR*	Serine recombinase (Putative DNA-invertase from lambdoid prophage Rac)	9.9			

Transcript id	Gene	Protein/Function	Fold-change (24h/1.5 OD) in WT		Fold-change (24h/1.5 OD) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC76042-1	*exbB*	Biopolymer transport protein	9.8	5.9	5.3	
AAC74618-1	*pinQ*	Serine recombinase (Putative DNA-invertase from lambdoid prophage Qin)	9.8	10.1		
AAC74617-1	*ydfK*	Cold shock protein	9.8			
AAC73143-1	*carA*	Carbamoyl-phosphate synthase small chain	9.6			
AAC76362-1	*bfd*	Bacterioferritin-associated ferredoxin	9.3	11.1	5.3	
EBE00001515050	*ryhB*	Small regulatory RNA	9.3			
AAC74457-1	*ynaE*	Uncharacterized protein	9.3			
AAC73685-1	*fepA*	Ferrienterobactin receptor	9.2		18.5	20.2
AAC77038-1	*yjcH*	Inner membrane protein	8.9	8.6	6.4	
AAC77277-1	*gntP*	High-affinity gluconate transporter	8.7	9.1		
AAC76625-1	*mtlR*	Mannitol operon repressor	8.7			
AAC75723-1	*nrdF*	Ribonucleoside-diphosphate reductase 2 subunit beta	8.4			
AAC76187	*yhbO*	Protein deglycase 2	8.4			
AAC73610-1	*hyi*	Hydroxypyruvate isomerase	8.3		4.2	
AAC76118-1	*higB*	mRNA interferase toxin	8.3			
AAC74387-1	*pspB*	Phage shock protein B	8.3			
AAT48230-1	*fadA*	3-ketoacyl-CoA thiolase (Fatty acid oxidation complex subunit beta)	8.2		22.3	5.9
AAC73893-1	*mcbA*	MqsR-controlled colanic acid and biofilm protein A	8.2			
AAC76451-1	*glpD*	Aerobic glycerol-3-phosphate dehydrogenase	8.1		6.4	
AAC74245-1	*ycgX*	Uncharacterized protein	8.0		4.2	17.6
AAC76105-1	*yqjH*	NADPH-dependent ferric-chelate reductase	8.0		13.4	
AAC75851-1	*ygdI*	Uncharacterized lipoprotein	7.9			

Transcript id	Gene	Protein/Function	Fold-change (24h/1.5 OD) in WT		Fold-change (24h/1.5 OD) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC76628-1	*lldR*	Putative L-lactate dehydrogenase operon regulatory protein	7.8	4.6	9.1	4.6
AAC77017-1	*yjbL*	Uncharacterized protein	7.7		21.6	
AAC73261-1	*fhuA*	Ferrichrome-iron receptor	7.6		7.8	32.9
AAC76627-1	*lldP*	L-lactate permease	7.6	5.7	13.8	11.6
AAC74505-1	*ydcJ*	Uncharacterized protein	7.6	7.7		
AAC75370-1	*argT*	Lysine/arginine/ornithine-binding periplasmic protein	7.4	7.9	5.6	8.3
AAC73698-1	*entH*	Proofreading thioesterase EntH	7.4	13.3	11.9	8.9
AAD13441-1	*yddM*	HTH-type transcriptional regulator	7.4	9.5		
AAC74186-1	*fhuE*	Ferric coprogen/ferric rhodotorulic acid outer membrane transporter	7.3		8.0	21.2
AAC74123-1	*csgE*	Curli production assembly/transport component	7.3	7.8		
AAC75399-1	*yfcV*	Uncharacterized fimbrial-like protein	7.1		16.3	4.2
AAC74193-1	*ndh*	NADH dehydrogenase	7.1	6.8		
AAC76495-1	*tusA*	Sulfur carrier protein	7.1			
AAC73693-1	*fepB*	Ferrienterobactin-binding periplasmic protein	7.0		11.5	15.5
AAC76041-1	*exbD*	Biopolymer transport protein	6.9		4.3	
AAC76849-1	*fadB*	Fatty acid oxidation complex subunit alpha	6.9	4.1	26.9	12.9
AAC75877-1	*lysA*	Diaminopimelate decarboxylase	6.9			
AAC75717-1	*alaE*	L-alanine exporter	6.8	7.6	13.6	
EBE00001514864	*efeU*	pseudogene	6.8		6.5	13.5
AAC75892-1	*ygeI*	Uncharacterized protein	6.8		5.0	30.4
AAC75020-1	*yodD*	Uncharacterized protein	6.8			
AAC76478-1	*ugpB*	sn-glycerol-3-phosphate-binding periplasmic protein	6.7		4.2	

Transcript id	Gene	Protein/Function	Fold-change (24h/1.5 OD) in WT		Fold-change (24h/1.5 OD) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC75400-1	*sixA*	Phosphohistidine phosphatase	6.7	6.1		
ACO59996-1	*yobI*	Uncharacterized protein	6.6		7.1	
AAC74389	*pspD*	Phage shock protein D	6.6			
AAC76611-1	*yiaW*	Inner membrane protein	6.4		14.1	
AAC75888-1	*yqeK*	Uncharacterized protein	6.4		20.8	41.5
AAD13440-1	*fdnI*	Formate dehydrogenase	6.4			
EBE00001515078	*ssrS*	6S RNA	6.3		19.6	
AAC76079-1	*ygiL*	Uncharacterized fimbrial-like protein	6.3		4.1	
AAC77032-1	*soxS*	Regulatory protein	6.3			
AAC73768-1	*asnB*	Asparagine synthetase B	6.2	4.5		
EBE00001515127	*thrW*	tRNA	6.2			
AAC77015-1	*yjbJ*	UPF0337 protein	6.2			
AAC73691-1	*fepD*	Ferric enterobactin transport system permease protein	6.1		11.4	
AAC74754-1	*sufA*	Iron-sulfur cluster insertion protein	6.1		22.6	33.8
AAC73861-1	*bioA*	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase	6.0			
AAC73433-1	*prpR*	Propionate catabolism operon regulatory protein	6.0	8.8		
AAC74386-1	*pspA*	Phage shock protein A	6.0			
AAC73692-1	*entS*	Enterobactin exporter	5.9		11.9	
EBE00001514922	*ygaQ*	pseudogene	5.9		24.2	39.9
AAC73862-1	*bioB*	Biotin synthase	5.9	11.9		
AAC74753-1	*sufB*	FeS cluster assembly protein	5.8		20.0	27.5
AAC74534-1	*yncE*	Uncharacterized protein	5.8		11.6	11.3
EBE00001514849	*yhdW*	pseudogene	5.8			
EBE00001515049	*dicF*	Small regulatory RNA	5.7		6.1	
AAC76108-1	*patA*	Putrescine aminotransferase	5.7			
AAC76678-1	*xanP*	Xanthine permease	5.7			
AAC77030-1	*yjcB*	Uncharacterized protein	5.6			
AAC77037-1	*actP*	Cation/acetate symporter	5.5		4.1	

Transcript id	Gene	Protein/Function	Fold-change (24h/1.5 OD) in WT		Fold-change (24h/1.5 OD) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC73609-1	*gcl*	Glyoxylate carboligase	5.5			
AAC74379-1	*puuA*	Gamma-glutamylputrescine synthetase	5.4	5.4	18.0	12.2
EBE00001515099	*ssrA*	tmRNA	5.4		5.7	
AAC74122-1	*csgF*	Curli production assembly/transport component	5.4			
AAC73325-1	*fadE*	Acyl-coenzyme A dehydrogenase	5.3		22.6	
AAC77321-1	*yjjQ*	Putative transcription factor	5.3		10.3	
AAC74896-1	*mgrB*	PhoP/PhoQ regulator	5.3			
AAC73659-1	*ybcV*	Uncharacterized protein	5.2		11.6	
AAC73330-1	*dinJ*	Antitoxin	5.2			
AAC75632-1	*grcA*	Autonomous glycyl radical cofactor	5.2		-19.0	-4.1
AAC76629-1	*lldD*	L-lactate dehydrogenase	5.2	4.8		
AAC74124-1	*csgD*	CsgBAC operon transcriptional regulatory protein	5.1	7.8		
AAC76425-1	*hslR*	Heat shock protein 15	5.1			
EBE00001515040	*leuV*	tRNA	5.1			
AAC77314-1	*lgoD*	L-galactonate-5-dehydrogenase	5.1			
AAC75437-1	*lpxP*	Lipid A biosynthesis palmitoleoyltransferase	5.1			
EBE00001515030	*proK*	tRNA	5.1			
AAC76077-1	*ribB*	3,4-dihydroxy-2-butanone 4-phosphate synthase	5.1	7.0		
AAC74682-1	*tus*	DNA replication terminus site-binding protein	5.1			
AAC76751-1	*pstS*	Phosphate-binding protein	5.0			
AAC77269-1	*fimE*	Type 1 fimbriae regulatory protein	4.9	7.5	11.4	8.2
EBE00001514948	*ydfJ*	pseudogene	4.9		5.7	9.9
EBE00001515105	*glmY*	Small regulatory RNA	4.9			
AAC77068-1	*phnB*	Conserved protein	4.9			

Transcript id	Gene	Protein/Function	Fold-change (24h/1.5 OD) in WT		Fold-change (24h/1.5 OD) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC75713-1	*yqaE*	UPF0057 membrane protein	4.9			
EBE00001514796	*yddK*	Leucine-rich repeat domain-containing protein	4.8		25.5	
AAC76821-1	*yigG*	Inner membrane protein	4.8		7.1	
AAC76579-1	*yiaG*	HTH-type transcriptional regulator	4.8			
ABD18681-1	*yoeB*	Toxin	4.8			
AAC73450-1	*mhpA*	3-(3-hydroxy-phenyl) propionate/3-hydroxycinnamic acid hydroxylase	4.7		11.5	
AAC73438-1	*prpE*	Propionate--CoA ligase	4.7		7.2	
ACO60001-1	*yohP*	Uncharacterized membrane protein	4.7		13.4	
AAC73224-1	*pdhR*	Pyruvate dehydrogenase complex repressor	4.7	4.3		
AAC74943-1	*torY*	Cytochrome c-type protein	4.7			
AAC76549-1	*yhjG*	AsmA family protein	4.7	7.4		
ADO17949-1	*mgtL*	Regulatory leader peptide for mgtA	4.6		5.5	
AAC74554-1	*bdm*	Biofilm-dependent modulation protein	4.6			
EBE00001514974	*dsrA*	Small regulatory RNA	4.6			
AAC73941-1	*potF*	Putrescine-binding periplasmic protein	4.5	4.7		
EBE00001514784	*ybbD*	pseudogene	4.4		13.4	4.5
AAC73538-1	*bolA*	DNA-binding transcriptional regulator	4.4			
AAC75019-1	*dsrB*	Protein DsrB	4.4	8.9		
AAC74635-1	*hokD*	Protein HokD	4.4	5.0		
AAC77067-1	*phnC*	Phosphonates import ATP-binding protein	4.3		15.7	
AAC74105-1	*phoH*	Phosphate starvation-inducible protein	4.3			
EBE00001515009	*ryeA*	Small RNA	4.3			
AAC75196-1	*yohC*	Inner membrane protein	4.3	4.5		
AAC73553-1	*glnK*	Nitrogen regulatory protein P-II 2	4.2		6.9	

Transcript id	Gene	Protein/Function	Fold-change (24h/1.5 OD) in WT		Fold-change (24h/1.5 OD) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
EBE00001514875	*yicT*	pseudogene	4.2		4.1	
AAC76396-1	*frlB*	Fructoselysine 6-phosphate deglycase	4.2			
AAC73337-1	*yafO*	mRNA interferase toxin	4.2	4.9		
AAC73329-1	*yafQ*	mRNA interferase toxin	4.2			
AAC75558-1	*yfgH*	Uncharacterized lipoprotein	4.2			
AAC74752	*sufC*	Probable ATP-dependent transporter	4.1		16.7	12.4
AAC74720-1	*ydhL*	Uncharacterized protein	4.1			
AAC74668-1	*ynfM*	Inner membrane transport protein	4.1			
AAC75202-1	*yohJ*	UPF0299 membrane protein	4.1			
AAC73684-1	*entD*	Enterobactin synthase component D	4.0		11.5	
AAC74629-1	*essQ*	Prophage lysis protein S homolog	4.0		9.0	

(The transcripts also present in microarray are mention with fold-change value. The transcripts also showed higher abundance in Δ*rpoS* were RpoS-independent and fold-change are mention. The transcripts showed low abundance in Δ*rpoS* were RpoS-dependent and fold-change are mention with negative (-) sign. All the transcripts were with fold-change ≥ 4 and FDR adjusted *p*-value ≤ 0.05).

APPENDIX 6: Complete list of transcripts that were higher in abundance during prolonged-incubation (48h relative to 24h of incubation)

Table 14: Transcripts that are higher in abundance during 48h relative to 24h of incubation.

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC73467-1	*yaiS*	Uncharacterized deacetylase	78.7	66.4	-4.3	
AAC76154-1	*tdcR*	Threonine dehydratase operon activator protein	73.5		-9.5	
AAC74575-1	*ydeQ*	Uncharacterized fimbrial-like protein	70.6		-14.5	-13.1
AAC73251-1	*yadV*	Probable fimbrial chaperone	67.3	62.8	-5.5	
AAC73633-1	*sfmC*	Probable fimbrial chaperone	62.6	16.7	-6.5	
AAC73404-1	*rclC*	Inner membrane protein	61.8	88.8	-4.3	-5.6
AAC73250-1	*htrE*	Outer membrane usher protein	61.8	75.3	-10.5	
AAC73635-1	*sfmH*	Uncharacterized fimbrial-like protein	61.8	59.5	-7.6	
EBE00001514858	*ybfQ*	pseudogene	60.9		-5.8	
AAC75399-1	*yfcV*	Uncharacterized fimbrial-like protein	60.7	61.4	-4.5	
EBE00001514874	*yddL*	pseudogene	59.5	110.7	-12.0	-21.1
AAC75816-1	*ygcW*	Uncharacterized oxidoreductase	58.6	4.7	-8.8	
AAC73406-1	*rclB*	Reactive chlorine resistance protein B	56.7	54.8		
AAC77267-1	*nanC*	N-acetylneuraminic acid outer membrane channel protein	55.9	30.8	-8.3	-14.8
AAC73252-1	*yadN*	Uncharacterized fimbrial-like protein	55.3	84.4	-16.2	
AAC76156-1	*yhaC*	Uncharacterized protein	54.8	42.6	-7.9	-37.6
AAC76247-1	*yhcA*	Uncharacterized fimbrial chaperone	54.6	82.8	-4.1	-10.6
AAC77297-1	*yjiS*	Uncharacterized protein	54.6		-4.6	
ABP93436-1	*ylcI*	Uncharacterized protein	53.9			
AAC74569-1	*yddA*	Inner membrane ABC transporter ATP-binding protein	53.5	15.6	-19.4	-28.2
EBE00001514876	*ypjC*	pseudogene	51.8	30.2	-19.4	-28.2

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
EBE00001514796	*yddK*	pseudogene	51.2		-4.0	
AAC77036-1	*yjcF*	Uncharacterized protein	51.2		-5.4	
AAC75146-1	*yegR*	Uncharacterized protein	51.1	13.0		
EBE00001514920	*ycgH-1*	pseudogene	49.9			
AAC73704-1	*ybdO*	HTH-type transcriptional regulator	48.9	39.7		
EBE00001514759	*yedN*	pseudogene	48.6		-5.1	
EBE00001514780	*ydeT*	pseudogene	47.4	44.3	-6.7	
EBE00001514757	*ybfL*	pseudogene	46.3		-15.7	
AAC73249-1	*yadM*	Uncharacterized fimbrial-like protein	46.2	30.0	-8.7	-10.4
AAC77138-1	*yjfI*	Uncharacterized protein	46.0		-5.9	
AAC77210-1	*yjgL*	Uncharacterized protein	45.3	38.8		
AAC74573-1	*safA*	Two-component-system connector protein	45.0		-7.6	
AAC73660-1	*ybcW*	Uncharacterized protein	44.9			
EBE00001514872	*yhiS*	pseudogene	44.3	52.3		
EBE00001514766	*yeeL*	pseudogene	44.0	27.4	-8.5	
AAC73856-1	*ybhH*	Putative isomerase	43.9		-15.9	
AAC77301-1	*mcrC*	5-methylcytosine-specific restriction enzyme subunit	42.4			
AAC74491-1	*ynbB*	Uncharacterized protein	42.3		-8.6	-16.6
AAC74110-1	*ycdT*	Probable diguanylate cyclase	41.7	71.6		
EBE00001514958	*yoeA*	pseudogene	40.4	67.7		
AAC74570-1	*ydeM*	Anaerobic sulfatase-maturating enzyme homolog	40.3	10.5	-5.3	
EBE00001514880	*yrhA*	pseudogene	40.3	17.4		
AAC74576-1	*ydeR*	Uncharacterized fimbrial-like protein	40.2		-9.8	-13.9
EBE00001514927	*yfdM*	pseudogene	40.0	31.5	-9.6	

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
EBE00001514828	*ykfJ*	pseudogene	40.0	60.3		
AAC76620-1	*yibG*	Uncharacterized protein	39.7		-6.1	-30.9
AAC77302-1	*mcrB*	5-methylcytosine-specific restriction enzyme B	39.6	28.1		
EBE00001514755	*yoeG*	pseudogene	39.5			
AAC76086-1	*yqiJ*	Inner membrane protein	39.1	24.6		
EBE00001514789	*yncI*	pseudogene	39.1	24.5	-5.8	
AAC76587-1	*yiaB*	Inner membrane protein	38.5	64.4	-7.0	-15.5
EBE00001514754	*yfdL*	pseudogene	38.1	50.8	-15.2	
AAC74542-1	*ydcC*	H repeat-associated putative transposase	37.7		-12.1	
EBE00001514905	*yejO*	pseudogene	37.7	27.6		
AAC75434-1	*yfdX*	Protein YfdX	36.9	60.2		
EBE00001514830	*ygeK*	pseudogene	36.9	19.7	-12.0	-7.1
AAC73561-1	*maa*	Maltose O-acetyltransferase	36.8			
EBE00001514854	*yhcE*	pseudogene	36.6	4.3		
ACO60004-1	*yqcG*	Uncharacterized protein	36.4		-4.7	
AAC76155-1	*yhaB*	Uncharacterized protein	36.1	41.0		
AAC77044-1	*yjcS*	Putative alkyl/aryl-sulfatase	36.0	27.1	-11.5	
AAC74761-1	*ydiN*	Inner membrane transport protein	35.9		-17.8	
AAC75890-1	*ygeG*	Uncharacterized protein	35.7		-7.6	-4.8
AAC76508-1	*yhhH*	Uncharacterized protein	35.6	42.0		
EBE00001514801	*ybcY*	pseudogene	35.4	62.0		
AAC75971-1	*cmtB*	Mannitol-specific cryptic phosphotransferase enzyme IIA component	35.2			
EBE00001514832	*yneL*	pseudogene	34.8	67.5		
AAC76467-1	*yhhZ*	Uncharacterized protein	34.8	45.0	-8.6	
AAD13451-1	*yiiE*	Uncharacterized protein	34.8		-8.8	-27.8

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
EBE00001514937	*yibU*	pseudogene	34.6		-5.4	
EBE00001514844	*yhiL*	pseudogene	34.5	21.8	-7.9	
AAC75171-1	*yehC*	Probable fimbrial chaperone	34.3		-7.5	
AAC74568-1	*yddB*	Uncharacterized protein	34.3	16.6		
AAC76273-1	*aaeA*	p-hydroxybenzoic acid efflux pump subunit	33.9	26.7		
AAC73381-1	*yagL*	Uncharacterized protein	33.9	56.7		
AAC73636-1	*sfmF*	Uncharacterized fimbrial-like protein	33.8			
AAC75170-1	*yehB*	Outer membrane usher protein	33.6		-11.7	-44.5
EBE00001514774	*mdtQ*	pseudogene	33.6		-7.7	
AAC76872-1	*ompL*	Porin	33.5	36.2	-8.4	
AAC74490-1	*ynbA*	Inner membrane protein	33.4	19.4	-7.0	
EBE00001514771	*ygeN*	pseudogene	33.1	9.0	-5.8	
AAC75423-1	*dsdC*	HTH-type transcriptional regulator	32.8	66.0		
EBE00001514821	*yibW*	pseudogene	32.5			
AAC73749-1	*ybeU*	Uncharacterized protein	32.5		-4.9	-22.5
AAC75426-1	*emrY*	Probable multidrug resistance protein	32.4	43.1	-19.2	-11.7
EBE00001515027	*rseX*	pseudogene	32.2			
AAC75416-1	*yfdN*	Uncharacterized protein	31.7		-5.5	-35.0
AAC74572-1	*ydeO*	HTH-type transcriptional regulator	31.6	33.5	-6.0	-44.6
AAC75885-1	*yqeH*	Uncharacterized protein	31.5	47.4	-4.9	-22.9
AAC76859-1	*yihF*	Uncharacterized protein	31.3	70.2	-19.9	-20.1
AAC76171-1	*kbaY*	D-tagatose-1,6-bisphosphate aldolase subunit	31.1	9.6		
EBE00001514834	*yjhR*	pseudogene	30.8	51.6		
AAC76743-1	*bglH*	Cryptic outer membrane porin	30.6		-6.9	-31.9
b1157-1	*stfE*	pseudogene	30.6	48.0	-4.6	

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC76509-1	*yhhI*	H repeat-associated putative transposase	30.4			
AAC73445-1	*lacA*	Galactoside O-acetyltransferase	30.1	8.4		
AAC75413-1	*yfdK*	Uncharacterized protein	30.1	36.5		
AAC74239-1	*ymfS*	Uncharacterized protein	30.1		-6.6	
AAC73247-1	*yadK*	Uncharacterized fimbrial-like protein	29.9		-8.6	-20.7
AAC76274-1	*aaeX*	DUF1656 domain-containing protein	29.8			
EBE00001514773	*oweS*	pseudogene	29.8		-7.9	-32.2
AAC77139-1	*yjfJ*	Uncharacterized protein	29.6			
AAC73446-1	*lacY*	Lactose permease	29.5			
AAC77268-1	*fimB*	Type 1 fimbriae regulatory protein	29.5	23.4		
AAC74982-1	*yecF*	Uncharacterized protein	29.4	54.4		
EBE00001514894	*pinH*	pseudogene	29.3	54.3	-5.2	
EBE00001514922	*ygaQ*	pseudogene	29.3		-13.2	
AAC73747-1	*djlB*	Uncharacterized J domain-containing protein	29.2	14.7	-4.2	
EBE00001514791	*yibV*	pseudogene	29.1		-9.7	-19.3
AAC74652-1	*rspB*	Starvation-sensing protein	29.1	26.9		
AAC73634-1	*sfmD*	Outer membrane usher protein	29.0		-5.9	
AAC75427-1	*emrK*	Probable multidrug resistance protein	28.7	40.2	-4.9	-28.7
AAC75716-1	*stpA*	DNA-binding protein StpA	28.4	69.8		
AAC74108-1	*pgaB*	Poly-beta-1,6-N-acetyl-D-glucosamine N-deacetylase	28.3			
EBE00001514837	*yjgX*	pseudogene	27.6	31.8		
AAC75891-1	*ygeH*	Uncharacterized protein	27.6		-13.9	
ABD18639-1	*ylcG*	Uncharacterized protein	27.5		-14.7	
AAC75034-1	*yedV*	Probable sensor-like histidine kinase	27.3	26.6		
AAC73466-1	*yaiP*	Uncharacterized glycosyltransferase	27.3			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC76479-1	*livF*	High-affinity branched-chain amino acid transport ATP-binding protein	27.1	44.3		
AAC76150-1	*tdcD*	Propionate kinase	27.0			
AAC75431-1	*yfdV*	Uncharacterized transporter	27.0	48.7	-14.5	
AAC77141-1	*yjfL*	UPF0719 inner membrane protein	27.0		-8.2	-21.4
AAC74983-1	*sdiA*	Regulatory protein	26.6	36.4		
AAC76682-1	*setC*	Sugar efflux transporter C	26.5		-6.1	
AAC74114-1	*ycdU*	Uncharacterized protein	26.5			
AAC76871-1	*yihN*	Inner membrane protein	26.3		-5.0	
AAC73746-1	*ybeR*	Uncharacterized protein	26.3		-8.3	
EBE00001515010	*pawZ*	pseudogene	26.1			
AAC74744-1	*ydhY*	Uncharacterized ferredoxin-like protein	26.1			
EBE00001514839	*yoeH*	pseudogene	25.9			
AAC77093-1	*cadB*	Probable cadaverine/lysine antiporter	25.7	9.9	-6.5	
AAC75714-1	*ygaV*	Probable HTH-type transcriptional regulator	25.7			
EBE00001514947	*ydfJ*	pseudogene	25.7	21.3		
AAC74086-1	*yccE*	Uncharacterized protein	25.6			
AAC73420-1	*yahC*	Uncharacterized protein	25.5		-7.8	-41.5
AAC75132-1	*yegJ*	Uncharacterized protein	25.5	10.9		
AAC75333-1	*yfbN*	Uncharacterized protein	25.3	30.0	-5.7	
AAC75240-1	*yejE*	Inner membrane ABC transporter permease protein	24.9	65.5		
AAC73813-1	*ybgD*	Uncharacterized fimbrial-like protein	24.9		-7.5	
AAC74222-1	*ymfE*	Uncharacterized protein	24.8	16.7		
AAC76177-1	*yraI*	Probable fimbrial chaperone	24.7	21.1		
EBE00001514955	*ycjV*	pseudogene	24.5		-6.8	
AAC73600-1	*ybbC*	Uncharacterized protein	24.5	27.2		
AAC74111-1	*insF*	Transposase InsF for insertion sequence IS3D	24.4			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
EBE00001514846	*eaeH*	pseudogene	24.3	11.7	-10.3	-41.5
EBE00001514823	*ykgQ*	pseudogene	24.3		-16.9	-17.7
AAC73397-1	*ecpR*	HTH-type transcriptional regulator	24.2			
AAC77255-1	*yjhI*	Uncharacterized HTH-type transcriptional regulator	24.2	27.3		
AAC75715-1	*ygaP*	Inner membrane protein	24.1	7.7		
AAC75888-1	*yqeK*	Uncharacterized protein	24.0		-7.7	-10.4
AAC76246-1	*gltF*	Periplasmic protein	23.9			
AAC74577-1	*ydeS*	Uncharacterized fimbrial-like protein	23.7	34.0	-4.3	-10.1
AAC73559-1	*ylaB*	Putative cyclic-di-GMP phosphodiesterase	23.4	26.2		
AAC74252-1	*ycgG*	Uncharacterized protein	23.4	79.5	-5.4	-11.2
AAC73857-1	*ybhI*	Inner membrane protein	23.2			
AAC73248-1	*yadL*	Uncharacterized fimbrial-like protein	22.9	40.4	-4.6	
AAC74238-1	*stfP*	Uncharacterized protein StfP from lambdoid prophage e14 region	22.9	22.8	-5.3	
AAC75430-1	*yfdE*	Acetyl-CoA:oxalate CoA-transferase	22.8	39.2	-5.5	
AAC73550-1	*decR*	DNA-binding transcriptional activator	22.8			
EBE00001514868	*yjbI*	pseudogene	22.8		-6.8	
AAC76257-1	*nanA*	N-acetylneuraminate lyase	22.7	31.4	5.8	8.2
EBE00001514761	*rhsJ*	pseudogene	22.5	37.4		
AAC77140-1	*yjfK*	Uncharacterized protein	22.3		-12.6	-42.5
AAC73748-1	*ybeT*	Sel1-repeat-containing protein	22.2		-8.8	-46.2
EBE00001514814	*ins1*	pseudogene	22.0			
AAC73562-1	*hha*	Hemolysin expression-modulating protein	21.9		4.1	4.0
EBE00001514795	*pbl*	pseudogene	21.9		-12.5	
AAC74764-1	*ydiF*	Acetate CoA-transferase	21.7	57.1	-6.6	

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC73380-1	*yagK*	Uncharacterized protein	21.6			
AAC76586-1	*yiaA*	Inner membrane protein	21.3	56.3		
AAC75432-1	*oxc*	Oxalyl-CoA decarboxylase	21.3		-5.8	-50.1
EBE00001514812	*yrhC*	pseudogene	21.3		-4.4	
AAC74244-1	*iraM*	Anti-adapter protein	21.2	28.6	-5.2	
EBE00001514772	*ymdE*	pseudogene	20.6	32.6	-7.9	-22.0
AAC75239-1	*yejB*	Inner membrane ABC transporter permease protein	20.2	16.9		
AAC73390-1	*yagU*	Inner membrane protein	20.2	32.1		
AAC76618-1	*yibA*	Putative lyase containing HEAT-repeat	20.2	20.1		
AAC76609-1	*yiaU*	Uncharacterized HTH-type transcriptional regulator	20.2	26.6		
AAC75035-1	*yedW*	Probable transcriptional regulatory protein	20.1	16.8		
AAC74619-1	*tfaQ*	Tail fiber assembly protein homolog from lambdoid prophage Qin	20.1	11.5		
AAT48246-1	*yjgN*	Inner membrane protein	19.9		-10.5	-14.0
AAC74455-1	*tfaR*	Tail fiber assembly protein homolog from lambdoid prophage Rac	19.9			
AAC74071-1	*gfcB*	Uncharacterized lipoprotein	19.9	11.0	-5.7	
AAC76394-1	*yhfL*	Uncharacterized protein	19.9	44.9		
ABP93440-1	*ymgI*	Uncharacterized protein	19.9		-12.3	
AAC74620-1	*stfQ*	Side tail fiber protein homolog from lambdoid prophage Qin	19.8	15.3		
AAC77303-1	*symE*	Toxic protein	19.8	28.1		
AAC74760-1	*ydiM*	Inner membrane transport protein	19.7		-7.0	
AAC74242-1	*pinE*	Serine recombinase	19.7	27.5	-5.3	-9.1
AAC75801-1	*casB*	CRISPR system Cascade subunit	19.6			
AAT48165-1	*ygjI*	Inner membrane transport protein	19.6		-4.9	
AAC75926-1	*uacT*	Uric acid transporter	19.6			
AAC74221-1	*ymfD*	Uncharacterized protein	19.5			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC75241-1	*yejF*	Uncharacterized ABC transporter ATP-binding protein	19.4	14.3		
AAC75420-1	*yfdR*	Uncharacterized protein	19.4		-8.2	
AAC74855-1	*yeaI*	Inner membrane protein	19.3	39.9	-10.2	-24.2
AAC74240-1	*tfaE*	Tail fiber assembly protein homolog from lambdoid prophage e14	19.3	32.6		
AAC73650-1	*ybcO*	Uncharacterized protein	19.3			
AAC73673-1	*cusC*	Cation efflux system protein	19.2			
AAC74107-1	*pgaC*	Poly-beta-1,6-N-acetyl-D-glucosamine synthase	19.2	10.8	-5.3	-22.6
AAC74653-1	*rspA*	Starvation-sensing protein	19.0	35.5		
AAC77233-1	*yjgZ*	Uncharacterized protein	18.8	9.3		
EBE00001514778	*yehH*	pseudogene	18.7	11.3	-11.3	-17.1
AAC73468-1	*tauA*	Taurine-binding periplasmic protein	18.7		-6.7	
AAC73246-1	*yadC*	Uncharacterized fimbrial-like protein	18.7	43.0	-7.1	
AAC73382-1	*yagM*	Uncharacterized protein	18.7			
AAC76251-1	*yhcF*	Uncharacterized protein	18.7	19.2		
AAC73560-1	*ylaC*	Inner membrane protein	18.6	7.4		
AAC75172-1	*yehD*	Uncharacterized fimbrial-like protein	18.5			
AAC73338-1	*yafP*	Uncharacterized N-acetyltransferase	18.5	24.6		
AAC75433-1	*frc*	Formyl-CoA: oxalate CoA-transferase	18.4			
AAC74842-1	*ydjH*	Uncharacterized sugar kinase	18.4			
AAC73674-1	*cusF*	Cation efflux system protein	18.3	10.4		
EBE00001514969	*yaiX*	pseudogene	18.3		-9.6	
AAC75227-1	*psuK*	Pseudouridine kinase	18.2			
AAC75116-1	*wcaE*	Putative colanic acid biosynthesis glycosyl transferase	18.2		-4.2	-20.5
AAC76703-1	*yidL*	Uncharacterized HTH-type transcriptional regulator	18.2	27.2		

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC73551-1	*mdlA*	Multidrug resistance-like ATP-binding protein	18.1			
AAD13437-1	*yddG*	Aromatic amino acid exporter	18.0	14.3		
AAC75673-1	*yfjI*	Uncharacterized protein	18.0		-12.1	
AAC73637-1	*fimZ*	Fimbriae Z protein	17.8			
AAC76600-1	*yiaL*	DUF386 domain-containing protein	17.8	27.1	-11.3	
AAC76256-1	*nanT*	Sialic acid transporter	17.8	29.0	4.2	7.0
AAC75331-1	*yfbL*	Uncharacterized protein	17.8			
AAC74839-1	*ydjE*	Inner membrane metabolite transport protein	17.6			
AAC74856-1	*yeaJ*	Putative diguanylate cyclase	17.3			
AAC77281-1	*yjiC*	Uncharacterized protein	17.3			
AAC76272-1	*aaeB*	p-hydroxybenzoic acid efflux pump subunit	17.2	12.8		
AAC74600-1	*yneK*	Uncharacterized protein	17.1	33.8		
AAC76601-1	*yiaM*	2,3-diketo-L-gulonate TRAP transporter small permease protein	17.0		-4.7	-17.8
EBE00001514848	*ykgP*	pseudogene	17.0		-9.2	
EBE00001514817	*ycgI*	pseudogene	16.7			
AAC74401-1	*ompG*	Outer membrane protein G	16.6		-4.2	
EBE00001514906	*yiaM-1*	pseudogene	16.6			
AAC74567-1	*pqqL*	Probable zinc protease	16.5	8.5		
EBE00001514899	*yrdE*	pseudogene	16.5			
AAC75117-1	*wcaD*	Putative colanic acid polymerase	16.5		-7.6	-9.5
b0561-1	*tfaD*	pseudogene	16.4			
AAC76732-1	*tnaB*	Low affinity tryptophan permease	16.3	28.4	6.3	
AAC74261-1	*ycgJ*	Uncharacterized protein	16.3	11.4		
AAC75268-1	*napF*	Ferredoxin-type protein	16.2			
ABD18655-1	*ymiA*	Uncharacterized protein	16.2	11.0		

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC74229-1	*cohE*	Putative lambdoid prophage e14 repressor protein C2	16.1	10.9		
EBE00001514806	*ylbI*	pseudogene	16.0			
ABP93434-1	*ykfM*	Uncharacterized protein	16.0			
AAC74081-1	*torC*	Cytochrome c-type protein	15.9		-12.2	
AAC75047-1	*yeeN*	Probable transcriptional regulatory protein	15.9	23.0		
AAC75798-1	*casE*	CRISPR system Cascade subunit	15.8	23.7		
AAC74125-1	*csgB*	Minor curlin subunit	15.8		-8.5	-15.1
EBE00001514792	*tfaX*	pseudogene	15.8			
AAC75820-1	*ygcG*	UPF0603 protein	15.8	44.1	-9.6	-40.1
AAC73187-1	*leuO*	HTH-type transcriptional regulator	15.7			
AAC76087-1	*yqiK*	Inner membrane protein	15.7	26.3		
b2850-2	*ygeF*	pseudogene	15.6	30.5	-11.2	
ACO60005-1	*yqeL*	Uncharacterized protein	15.6			
AAC74571-1	*ydeN*	Uncharacterized sulfatase	15.5	12.0		
AAC74647-1	*dicB*	Division inhibition protein	15.4	21.7	-10.1	-6.1
AAC76255-1	*nanE*	Putative N-acetylmannosamine-6-phosphate 2-epimerase	15.4	12.8		
AAC77018-1	*yjbM*	Uncharacterized protein	15.1	8.9	-7.8	
AAC74255-1	*ymgD*	Uncharacterized protein	15.1	19.9		
AAC76248-1	*yhcD*	Uncharacterized outer membrane usher protein	15.0	10.7		
AAC75334-1	*yfbO*	Uncharacterized protein	15.0			
AAC75802-1	*casA*	CRISPR system Cascade subunit	14.9		-12.1	
AAC75332-1	*yfbM*	Protein YfbM	14.9			
EBE00001514798	*yoeF*	pseudogene	14.9	6.1		
AAC74245-1	*ycgX*	Uncharacterized protein	14.9	21.4		
AAC77356-1	*lasT*	Uncharacterized tRNA/rRNA methyltransferase	14.8			
AAC76642-1	*htrL*	Protein HtrL	14.7			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC73800-1	*ybfD*	H repeat-associated putative transposase	14.6		-7.0	
AAC77224-1	*idnD*	L-idonate 5-dehydrogenase (NAD(P)(+))	14.6			
EBE00001514865	*exoD*	pseudogene	14.6	5.4		
AAC75039-1	*zinT*	Metal-binding protein	14.5	17.0		
EBE00001514918	*ypdj*	pseudogene	14.5			
EBE00001514867	*intQ*	pseudogene	14.5	23.9		
AAC75774-1	*ygbA*	Uncharacterized protein	14.5	20.6		
ACO59992-1	*ynbG*	Uncharacterized protein	14.5			
EBE00001514882	*icdC*	pseudogene	14.2			
EBE00001514768	*ylbG*	pseudogene	14.2	30.4		
AAC76348-1	*gspA*	Putative general secretion pathway protein A	14.2	28.0	-4.9	
AAC73409-1	*ykgE*	Uncharacterized protein	14.2			
AAC75424-1	*dsdX*	D-serine transporter	14.1			
AAC76610-1	*yiaV*	Inner membrane protein	14.1	39.6		
ABD18651-1	*ymgF*	Inner membrane protein	14.1			
AAC75970-1	*cmtA*	PTS system mannitol-specific cryptic EIICB component	14.1	12.4	-4.0	
AAC73796-1	*ybfB*	Uncharacterized protein	14.1	17.8		
AAC75419-1	*yfdQ*	Uncharacterized protein	14.1			
AAC76740-1	*cbrC*	UPF0167 protein	14.1			
AAC73143-1	*carA*	Carbamoyl-phosphate synthase small chain	14.0	15.4		
EBE00001514775	*yohP*	pseudogene	14.0			
AAC74762-1	*ydiB*	Quinate/shikimate dehydrogenase	14.0			
AAC77094-1	*cadC*	Transcriptional activator	14.0			
AAC75605-1	*hmp*	Flavohemoprotein	13.9			
AAC76766-1	*asnC*	Regulatory protein	13.9	16.2		
AAC74515-1	*ydcO*	Inner membrane protein	13.8			
AAC77273-1	*fimD*	Outer membrane usher protein	13.8	4.1		

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC75421-1	*yfdS*	Uncharacterized protein	13.8		-4.6	
AAC76742-1	*yieL*	Uncharacterized protein	13.8			
ABP93441-1	*ymgJ*	Uncharacterized protein	13.8			
AAC73651-1	*rusA*	Crossover junction endodeoxyribonuclease	13.7	7.8	-6.1	-17.6
AAC74489-1	*ydbD*	Uncharacterized protein	13.7	24.9	-4.4	
AAC74024-1	*elfA*	Fimbrial subunit	13.5	9.1		
AAC74109-1	*pgaA*	Poly-beta-1,6-N-acetyl-D-glucosamine export protein	13.5	5.1		
AAC73652-1	*quuD*	Prophage antitermination protein Q homolog from lambdoid prophage DLP12	13.5		-11.8	
AAC75460-1	*xapA*	Purine nucleoside phosphorylase 2	13.5			
AAC77142-1	*yjfM*	Uncharacterized protein	13.5		-4.5	-18.8
AAC77235-1	*yjhB*	Putative metabolite transport protein	13.3			
AAC75169-1	*yehA*	Uncharacterized fimbrial-like protein	13.3	39.0	-11.9	
AAC74025-1	*elfD*	Probable fimbrial chaperone	13.2			
AAC75143-1	*ogrK*	Prophage late control protein	13.2			
AAC73659-1	*ybcV*	Uncharacterized protein	13.2		-6.1	
AAC73422-1	*yahE*	Uncharacterized protein	13.1		-6.2	-11.4
AAC73874-1	*ybhM*	Uncharacterized protein	13.1	17.0		
AAC76513-1	*yhiJ*	Uncharacterized protein	13.1	37.9		
AAC75824-1	*mazF*	Endoribonuclease toxin	13.0	19.1		
EBE00001514961	*yafU*	pseudogene	13.0	16.1	-4.1	
AAC73632-1	*sfmA*	Uncharacterized fimbrial-like protein	12.9	40.0	-6.2	
AAC73645-1	*ybcK*	Uncharacterized protein	12.9		-7.7	
AAC75515-1	*eutS*	Ethanolamine utilization protein	12.8			
AAC74266-1	*hlyE*	Hemolysin E	12.8	20.8		
AAC74954-1	*cheR*	Chemotaxis protein methyltransferase	12.7			
AAC74832-1	*ynjI*	Inner membrane protein	12.7	17.3		
AAC74841-1	*ydjG*	Uncharacterized oxidoreductase	12.7			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC77236-1	*yjhC*	Uncharacterized oxidoreductase	12.7	21.6		
AAC75405-1	*yfdF*	Uncharacterized protein	12.7			
AAC74243-1	*mcrA*	5-methylcytosine-specific restriction enzyme A	12.6	34.2		
AAC74275-1	*cvrA*	K(+)/H(+) antiporter NhaP2	12.6	12.8		
AAC74766-1	*ydiP*	Uncharacterized HTH-type transcriptional regulator	12.6	6.9		
EBE00001514816	*insc1*	pseudogene	12.5			
AAC74028-1	*ycbU*	Uncharacterized fimbrial-like protein	12.5			
AAC77250-1	*insA*	Insertion element IS1 7 protein	12.4			
ABP93458-1	*yjbS*	Uncharacterized protein	12.3		-7.7	
AAC75799-1	*casD*	CRISPR system Cascade subunit	12.2			
EBE00001514948	*ydfK*	pseudogene	12.2			
AAC76882-1	*frvA*	PTS system fructose-like EIIA component	12.2		-5.9	-14.0
AAC74682-1	*tus*	DNA replication terminus site-binding protein	12.1			
EBE00001514952	*ygaY*	pseudogene	12.1	15.8		
EBE00001515019	*ileY*	pseudogene	12.1			
AAC76021-1	*yghS*	Uncharacterized ATP-binding protein	12.1	8.8		
AAC76746-1	*bglG*	Cryptic beta-glucoside bgl operon antiterminator	12.0			
AAC75672-1	*alpA*	DNA-binding transcriptional activator	11.9			
AAC75017-1	*fliR*	Flagellar biosynthetic protein	11.9	6.0		
EBE00001514824	*ybfI*	pseudogene	11.9			
AAC73960-1	*hcp*	Hydroxylamine reductase	11.8			
EBE00001515014	*valV*	tRNA	11.8			
AAC73686-1	*fes*	Enterochelin esterase	11.7			
AAC75224-1	*yeiL*	Regulatory protein	11.6			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC76106-1	*yqjI*	Transcriptional regulator	11.6	8.7		
AAC76112-1	*ebgC*	Evolved beta-galactosidase subunit beta	11.5			
AAC74357-1	*cysB*	HTH-type transcriptional regulator	11.5	7.9		
AAC74223-1	*lit*	Bacteriophage T4 late gene expression-blocking protein	11.4	12.7		
AAC74157-1	*flgB*	Flagellar basal body rod protein FlgB (Putative proximal rod protein)	11.4			
AAC73152-1	*fixA*	Putative electron transfer flavoprotein	11.4		-5.2	
EBE00001514769	*afuB*	pseudogene	11.4			
EBE00001514765	*yjiT*	pseudogene	11.4			
AAC74648-1	*ydfD*	Uncharacterized protein	11.4	17.6	-5.5	
AAC76252-1	*yhcG*	Uncharacterized protein	11.4			
AAC76471-1	*yrhB*	Uncharacterized protein	11.4		-5.0	-7.5
AAC75242-1	*yejG*	Uncharacterized protein	11.3	8.7		
AAC74687-1	*uidC*	Membrane-associated protein	11.2	6.1	-5.0	-4.4
EBE00001514784	*ybbD*	pseudogene	11.2		-6.2	-21.0
AAC75435-1	*ypdI*	Uncharacterized lipoprotein	11.2	10.9		
AAC74843-1	*ydjI*	Uncharacterized protein	11.2			
AAC75329-1	*elaD*	Deubiquitinating enzyme	11.1		-4.3	
AAC76175-1	*agaI*	Putative galactosamine-6-phosphate isomerase	11.1		-6.7	
AAC74158-1	*flgC*	Flagellar basal body rod protein FlgB (Putative proximal rod protein)	11.0		-4.2	
AAC75386-1	*epmC*	Elongation factor P hydroxylase	10.9	24.7		
AAC75312-1	*ais*	Lipopolysaccharide core heptose(II)-phosphate phosphatase	10.9			
ACO59990-1	*ykgR*	Uncharacterized membrane protein	10.9			
AAC74314-1	*purU*	Formyltetrahydrofolate deformylase	10.8	7.9		

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC76152-1	*tdcB*	L-threonine dehydratase catabolic	10.8	5.2		
AAC74868-1	*leuE*	Leucine efflux protein	10.8	11.6		
AAC75920-1	*xanQ*	Xanthine permease	10.8	14.7		
AAC77089-1	*ghoS*	Endoribonuclease antitoxin	10.7		-7.2	-11.6
AAC74070-1	*gfcC*	Uncharacterized protein	10.7			
AAC74844-1	*ydjJ*	Uncharacterized zinc-type alcohol dehydrogenase-like protein	10.7			
EBE00001514840	*yaiT*	pseudogene	10.6	7.4		
EBE00001514962	*ykgA*	pseudogene	10.6	14.5		
AAC75040-1	*yodB*	Cytochrome b561 homolog 1	10.5	20.0		
AAC76739-1	*cbrB*	CreB-regulated gene B protein	10.4	16.8		
AAC75223-1	*rihB*	Pyrimidine-specific ribonucleoside hydrolase	10.4			
AAC73647-1	*ybcM*	Uncharacterized HTH-type transcriptional regulator	10.4	32.1		
AAC75422-1	*yfdT*	Uncharacterized protein	10.4			
AAC74045-1	*sxy*	Transcriptional coactivator for CRP	10.3			
AAT48233-1	*yihO*	**Putative sulfoquinovose importer**	10.3			
AAC75800-1	*casC*	CRISPR system Cascade subunit	10.2			
AAC73703-1	*ybdN*	Uncharacterized protein	10.2	34.4		
AAC75674-1	*yfjJ*	Uncharacterized protein	10.2			
AAC76744-1	*bglB*	6-phospho-beta-glucosidase	10.1		-5.3	
AAC73667-1	*envY*	Porin thermoregulatory protein	10.1			
AAC74224-1	*intE*	Prophage e14 integrase	10.0	10.9		
EBE00001514862	*insX*	pseudogene	10.0			
AAC73421-1	*yahD*	Putative ankyrin repeat protein	10.0		-5.9	-6.3
AAC73986-1	*ycaN*	Uncharacterized HTH-type transcriptional regulator	10.0			
AAC74743-1	*ydhV*	Uncharacterized oxidoreductase	10.0	7.7		

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC75882-1	*kduI*	4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase	9.9	7.8		
EBE00001514914	*rhsO*	pseudogene	9.9			
AAC73921-1	*yliF*	Putative lipoprotein	9.9			
AAC75914-1	*yqeC*	Uncharacterized protein	9.9			
AAC75653-1	*yfiN*	Probable diguanylate cyclase	9.8	9.5		
AAC75768-1	*hypA*	Hydrogenase 3 nickel incorporation protein	9.8	4.8		
AAC73407-1	*rclA*	Probable pyridine nucleotide-disulfide oxidoreductase	9.7	8.8	-5.9	
AAC74413-1	*insH*	Transposase InsH for insertion sequence element IS5F	9.7			
AAC73429-1	*cspE*	Uncharacterized protein	9.7		-4.8	
AAC76480-1	*livG*	High-affinity branched-chain amino acid transport ATP-binding protein	9.5	4.8		
AAC76821-1	*yigG*	Inner membrane protein	9.5	11.8		
AAC77320-1	*yjjP*	Inner membrane protein	9.5	5.8		
AAC74132-1	*mdoG*	Glucans biosynthesis protein G	9.4			
AAC77265-1	*nanS*	Probable 9-O-acetyl-N-acetylneuraminic acid deacetylase	9.4			
AAC75677-1	*yfjM*	Uncharacterized protein	9.4	13.0		
AAC77211-1	*argI*	Ornithine carbamoyltransferase subunit I	9.3	44.7		
AAC73705-1	*dsbG*	Thiol:disulfide interchange protein	9.3			
AAC76702-1	*yidK*	Uncharacterized symporter	9.3			
AAC74106-1	*pgaD*	Biofilm PGA synthesis protein	9.2	16.6		
AAC73180-1	*sgrR*	HTH-type transcriptional regulator	9.2			
ABD18695-1	*torI*	Response regulator inhibitor for tor operon	9.2			
AAC74943-1	*torY*	Cytochrome c-type protein	9.1	7.8		
AAC76431-1	*greB*	Transcription elongation factor	9.1	8.7		
AAC75949-1	*ygfA*	5-formyltetrahydrofolate cyclo-ligase	9.0			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC77279-1	*uxuB*	D-mannonate oxidoreductase	9.0	6.7		
AAC74029-1	*ycbV*	Uncharacterized fimbrial-like protein	9.0			
AAC75335-1	*yfbP*	Uncharacterized protein	9.0			
AAC73646-1	*ybcL*	UPF0098 protein	9.0	29.5		
AAC75015-1	*fliP*	Flagellar biosynthetic protein	8.9			
AAC73613-1	*ybbW*	Putative allantoin permease	8.9		-16.0	
AAC75695-1	*ypjA*	Uncharacterized outer membrane protein	8.9			
AAC75300-1	*glpT*	Glycerol-3-phosphate transporter	8.8			
AAT48186-1	*rhsB*	RhsB protein in rhs element	8.8	25.4	-4.4	-28.8
AAC76172-1	*agaB*	N-acetylgalactosamine-specific phosphotransferase enzyme IIB component 1	8.7		-6.5	
AAC74385-1	*pspF*	Psp operon transcriptional activator	8.7	4.7		
AAC76176-1	*yraH*	Uncharacterized fimbrial-like protein	8.7			
AAC76987-1	*arpA*	Ankyrin repeat protein A	8.6	23.3		
AAC77206-1	*bdcA*	Cyclic-di-GMP-binding biofilm dispersal mediator protein	8.6			
AAC73491-1	*aroL*	Shikimate kinase 2	8.6	8.8		
AAC73648-1	*ybcN*	Uncharacterized protein	8.6			
AAC74974-1	*yecR*	Uncharacterized protein	8.6			
AAC75803-1	*ygcB*	CRISPR-associated endonuclease/helicase	8.5			
AAC76585-1	*wecH*	O-acetyltransferase	8.5	4.7		
AAC73812-1	*ybgQ*	Uncharacterized outer membrane usher protein	8.5			
AAC73798-1	*ybfC*	Uncharacterized protein	8.5	11.5	-4.8	
ABD18656-1	*yciX*	Uncharacterized protein	8.5	11.1		
AAC76591-1	*xylG*	Xylose import ATP-binding protein	8.5			
AAC75652-1	*yfiR*	Uncharacterized protein	8.4	17.2		
AAC73825-1	*mngA*	PTS system 2-O-alpha-mannosyl-D-glycerate-specific EIIABC component	8.3		-9.4	
AAC75066-1	*cbtA*	Cytoskeleton-binding toxin	8.2			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC77278-1	*uxuA*	Mannonate dehydratase	8.2	14.0		
EBE00001514949	*rhsH*	pseudogene	8.2			
AAC76170-1	*agaS*	Putative tagatose-6-phosphate ketose/aldose isomerase	8.2			
AAC76179-1	*yraK*	Uncharacterized fimbrial-like protein	8.2	12.0		
AAC74128-1	*ymdA*	Uncharacterized protein	8.2			
AAC75396-1	*yfcS*	Probable fimbrial chaperone	8.1		-4.0	
AAC74072-1	*gfcA*	Threonine-rich inner membrane protein	8.1	8.4		
AAC75690-1	*yfjW*	Uncharacterized protein	8.1	17.9		
AAC76923-1	*metF*	5,10-methylenetetrahydrofolate reductase	8.0		8.7	
AAC73826-1	*mngB*	Mannosylglycerate hydrolase	8.0			
AAC73327-1	*yafJ*	Putative glutamine amidotransferase	8.0	9.8		
AAC74800-1	*ydjO*	Uncharacterized protein	8.0			
AAC76820-1	*yigF*	Uncharacterized protein	8.0			
AAT48194-1	*yiaN*	2,3-diketo-L-gulonate TRAP transporter large permease protein	7.9		-11.0	
AAC76527-1	*arsB*	Arsenical pump membrane protein	7.9	9.7		
AAC75167-1	*rcnA*	Nickel/cobalt efflux system	7.9	4.8		
AAT48195-1	*yiaY*	Probable alcohol dehydrogenase	7.9		-9.0	
EBE00001514815	*yqiG*	pseudogene	7.9	23.7		
EBE00001514820	*glvG*	pseudogene	7.9			
AAC77040-1	*nrfA*	Cytochrome c-552	7.8			
AAC77048-1	*alsA*	D-allose import ATP-binding protein	7.8			
AAC75267-1	*napD*	NapA signal peptide-binding chaperone	7.8			
AAC74642-1	*dicC*	Repressor protein of division inhibition gene	7.8		-10.6	
AAC75393-1	*yfcP*	Uncharacterized fimbrial-like protein	7.8	5.9		

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC73321-1	*yafT*	Uncharacterized lipoprotein	7.7			
AAC74716-1	*ydhJ*	Uncharacterized protein	7.7	4.9		
AAC74247-1	*bluF*	Blue light- and temperature-regulated antirepressor	7.6	11.6		
AAC73691-1	*fepD*	Ferric enterobactin transport system permease protein	7.6			
AAC75013-1	*fliN*	Flagellar motor switch protein	7.6			
AAC76649-1	*rfaY*	Lipopolysaccharide core heptose (II) kinase	7.6			
AAC77289-1	*yjiK*	Uncharacterized protein	7.6			
AAC76097-1	*ttdA*	L (+)-tartrate dehydratase subunit alpha	7.5			
AAC74459-1	*ompN*	Outer membrane protein N	7.5	7.1		
AAC73395-1	*ecpB*	Probable fimbrial chaperone	7.5			
EBE00001514909	*yjhZ*	pseudogene	7.5			
AAC73328-1	*yafK*	Putative L,D-transpeptidase	7.5			
AAC76731-1	*tnaA*	Tryptophanase	7.5	7.9		
AAT48200-1	*yidX*	Uncharacterized protein	7.5	22.4		
AAC74574-1	*ydeP*	Putative oxidoreductase	7.4			
AAC75746-1	*srlB*	PTS system glucitol/sorbitol-specific EIIA component	7.4		-4.0	
AAC75089-1	*ugd*	UDP-glucose 6-dehydrogenase	7.4			
AAC76083-1	*yqiH*	Uncharacterized fimbrial chaperone	7.4			
AAC76528-1	*arsC*	Arsenate reductase	7.3			
AAC73418-1	*yahA*	Cyclic di-GMP phosphodiesterase	7.3	15.8		
AAC75556-1	*yfgF*	Cyclic di-GMP phosphodiesterase	7.3			
AAC76543-1	*yhjA*	Probable cytochrome c peroxidase	7.3	11.5		
AAC74026-1	*elfC*	Probable outer membrane usher protein	7.3			
b0553-2	*nmpC*	pseudogene	7.3			
AAC74989-1	*fliA*	RNA polymerase sigma factor	7.3			
ABD18687-1	*yehK*	Uncharacterized protein	7.3			
AAC75755-1	*hydN*	Electron transport protein	7.2	10.1		
AAC75016-1	*fliQ*	Flagellar biosynthetic protein	7.2			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC73666-1	*ompT*	Outer membrane protein 3B	7.2			
AAC76608-1	*yiaT*	Putative outer membrane protein	7.2			
AAC75767-1	*hycA*	Formate hydrogenlyase regulatory protein	7.1		-4.1	
AAC76611-1	*yiaW*	Inner membrane protein	7.1	14.1	-9.2	
EBE00001514884	*yoeD*	pseudogene	7.1			
AAC77143-1	*yjfC*	Putative acid--amine ligase	7.1			
AAC76206-1	*yhbX*	Putative phosphoethanolamine transferase	7.1	5.7		
AAC76355-1	*gspI*	Putative type II secretion system protein I	7.1			
AAC73865-1	*bioD*	ATP-dependent dethiobiotin synthetase	7.0	14.1		
AAC74127-1	*csgC*	Curli assembly protein	7.0			
AAC77077-1	*adiY*	HTH-type transcriptional regulator	7.0		-10.0	
AAC74262-1	*pliG*	Inhibitor of g-type lysozyme	7.0	6.4		
AAC73961-1	*lysO*	Lysine exporter	7.0			
EBE00001514764	*gatR*	pseudogene	7.0	13.3	4.0	5.8
AAC75776-1	*pphB*	Serine/threonine-protein phosphatase 2	7.0			
AAC75651-1	*yfiL*	Uncharacterized protein	7.0	6.2		
AAC77266-1	*nanM*	N-acetylneuraminate epimerase	6.9			
AAC73649-1	*ninE*	Protein NinE homolog from lambdoid prophage DLP12	6.9		-6.4	
EBE00001514819	*glvC*	pseudogene	6.9			
EBE00001514877	*yeeW*	pseudogene	6.9	14.0		
AAC76878-1	*yiiG*	Uncharacterized protein	6.9			
AAC73688-1	*fepE*	Ferric enterobactin transport protein	6.8			
AAC76243-1	*yhcC*	Radical SAM family oxidoreductase	6.8			
AAC73987-1	*ycaK*	Uncharacterized NAD(P)H oxidoreductase	6.8			
AAD13452-1	*yiiF*	Uncharacterized protein	6.8			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC77272-1	*fimC*	Type 1 fimbriae periplasmic chaperone	6.7			
AAC76296-1	*envR*	Probable acrEF/envCD operon repressor	6.7			
AAC74267-1	*umuD*	DNA polymerase V protein	6.7	12.8		
AAC74391-1	*ycjM*	Putative sucrose phosphorylase	6.7			
AAC77051-1	*rpiB*	Ribose-5-phosphate isomerase B	6.7			
AAC74759-1	*ydiL*	Uncharacterized protein	6.7	15.5		
AAC75671-1	*yfjH*	Uncharacterized protein	6.7	17.3		
AAC73619-1	*allD*	Ureidoglycolate dehydrogenase	6.7			
AAC73149-1	*caiB*	L-carnitine CoA-transferase	6.6	13.7		
AAC73701-1	*ybdL*	Methionine aminotransferase	6.6			
AAC76392-1	*nirC*	Nitrite transporter	6.6	4.0		
AAC75524-1	*yffB*	Putative reductase	6.6	8.1		
AAC73750-1	*djlC*	Uncharacterized J domain-containing protein	6.6			
AAC73617-1	*allE*	(S)-ureidoglycine aminohydrolase	6.5			
AAC77004-1	*malE*	Maltose-binding periplasmic protein	6.5			
AAC73720-1	*dpiB*	Sensor histidine kinase	6.5			
AAC76151-1	*tdcC*	Threonine/serine transporter	6.5			
AAC76621-1	*yibH*	Inner membrane protein	6.4	9.0		
AAT48220-1	*rarD*	Putative transporter	6.4	19.8		
AAC75055-1	*insH1*	Transposase InsH for insertion sequence element IS5H	6.4			
AAC75511-1	*eutD*	Ethanolamine utilization protein	6.3			
AAD13438-1	*fdnG*	Formate dehydrogenase, nitrate-inducible, major subunit	6.3			
AAC76008-1	*pppA*	Leader peptidase	6.3	5.2		
AAC76173-1	*agaC*	N-acetylgalactosamine permease IIC component 1	6.3			
EBE00001514953	*yagP*	pseudogene	6.3			
AAC76349-1	*gspC*	Putative type II secretion system protein C	6.3			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC76358-1	*gspL*	Putative type II secretion system protein L	6.3			
AAC75668-1	*ratA*	Ribosome association toxin	6.3			
AAC76269-1	*argR*	Arginine repressor	6.2	5.0		
AAC73661-1	*nohD*	DNA-packaging protein NU1 homolog	6.2			
AAC74205-1	*ycfZ*	Inner membrane protein	6.2			
EBE00001514827	*ysaC*	pseudogene	6.2			
AAC74767-1	*ydiQ*	Putative electron transfer flavoprotein subunit	6.2			
AAC73156-1	*yaaU*	Putative metabolite transport protein	6.2		-10.4	
ABV59575-1	*ythA*	Uncharacterized protein	6.2			
AAC77285-1	*yjiG*	Inner membrane protein	6.1	10.5		
AAC75392-1	*yfcO*	Uncharacterized protein	6.1			
AAC73365-1	*fbpC*	Fe^{3+} ions import ATP-binding protein	6.0			
EBE00001514957	*lafU*	pseudogene	6.0			
ACO59991-1	*ymiB*	Putative protein	6.0			
AAC73598-1	*ybbP*	Uncharacterized ABC transporter permease	6.0	8.5		
AAC76022-1	*yghT*	Uncharacterized ATP-binding protein	6.0		-4.4	
AAC77161-1	*yjfZ*	Uncharacterized protein	6.0			
AAC76254-1	*nanK*	N-acetylmannosamine kinase	5.9			
AAT48170-1	*tdcE*	Keto-acid formate acetyltransferase	5.9			
AAC74765-1	*ydiO*	Probable acyl-CoA dehydrogenase	5.9			
AAC74105-1	*phoH*	Phosphate starvation-inducible protein	5.9			
EBE00001514807	*insO*	pseudogene	5.9			
EBE00001515138	*pauD*	pseudogene	5.9			
AAT48152-1	*ygeW*	Uncharacterized protein	5.9			
AAC77078-1	*adiA*	Biodegradative arginine decarboxylase	5.8			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC77049-1	*alsB*	D-allose-binding periplasmic protein	5.8			
AAC74991-1	*fliD*	Flagellar hook-associated protein 2	5.8			
AAC76063-1	*ygiZ*	Inner membrane protein	5.8		-4.2	-4.8
AAC74988-1	*fliZ*	Regulator of sigma S factor	5.8	5.4		
AAC74080-1	*torR*	TorCAD operon transcriptional regulatory protein	5.8	7.7		
AAC75394-1	*yfcQ*	Uncharacterized fimbrial-like protein	5.8			
AAC73419-1	*yahB*	Uncharacterized HTH-type transcriptional regulator	5.8	20.8		
AAC73810-1	*ybgO*	Uncharacterized protein	5.8			
ACO60006-1	*yqfG*	Uncharacterized protein	5.8			
AAC73675-1	*cusB*	Cation efflux system protein	5.7			
AAC73985-1	*ycaM*	Inner membrane transport protein	5.7	19.7		
EBE00001514889	*yfaH*	pseudogene	5.7			
AAC74456-1	*pinR*	Serine recombinase	5.7			
AAC76079-1	*ygiL*	Uncharacterized fimbrial-like protein	5.7			
AAC75238-1	*yejA*	Uncharacterized protein	5.7	4.8		
AAT48202-1	*yieK*	Uncharacterized protein	5.7			
AAC74871-1	*yeaV*	Uncharacterized transporter	5.7			
AAC75825-1	*mazE*	Antitoxin	5.6	10.0		
AAC75079-1	*hisL*	his operon leader peptide	5.6		-8.9	
AAC73337-1	*yafO*	mRNA interferase toxin	5.6	11.4		
AAC74079-1	*torT*	Periplasmic protein	5.6	16.5		
EBE00001514857	*prfH*	pseudogene	5.6			
EBE00001515032	*micC*	pseudogene	5.6			
AAC74840-1	*ydjF*	Uncharacterized HTH-type transcriptional regulator	5.6	20.2		
AAC74539-1	*ydcD*	Uncharacterized protein	5.6	4.3		
AAC75206-1	*yeiS*	Uncharacterized protein	5.6	20.5		
AAT48243-1	*ytfI*	Uncharacterized protein	5.6	35.5		
AAC73461-1	*yafD*	UPF0294 protein	5.6			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC75065-1	*cbeA*	Cytoskeleton bundling-enhancing antitoxin	5.5			
AAC74503-1	*trg*	Methyl-accepting chemotaxis protein III	5.5			
AAC73408-1	*rclR*	RCS-specific HTH-type transcriptional activator	5.5			
AAC74492-1	*ynbC*	Uncharacterized protein	5.5	5.5		
AAC75005-1	*fliF*	Flagellar M-ring protein	5.4		-4.5	
AAC73855-1	*ybhD*	Uncharacterized HTH-type transcriptional regulator	5.4			
AAC77251-1	*yjhU*	Uncharacterized transcriptional regulator	5.4			
AAC74256-1	*ymgG*	UPF0757 protein	5.4	11.3		
AAT48132-1	*blr*	Divisome-associated membrane protein	5.3			
AAC74516-1	*sutR*	HTH-type transcriptional regulator	5.3			
AAC76297-1	*acrE*	Multidrug export protein	5.3			
AAC73148-1	*caiC*	Probable crotonobetaine/carnitine-CoA ligase	5.3			
AAC76730-1	*tnaC*	Tryptophanase operon leader peptide	5.3			
AAC74634-1	*rem*	Uncharacterized protein	5.3			
AAC77286-1	*yjiH*	Uncharacterized protein	5.3	5.8		
AAC74410-1	*pgrR*	HTH-type transcriptional regulator	5.2	10.8		
AAC77006-1	*lamB*	Maltoporin	5.2			
AAC77005-1	*malK*	Maltose/maltodextrin import ATP-binding protein	5.2			
AAC75694-1	*ypjF*	Probable toxin	5.2			
AAC77274-1	*fimF*	Type 1 fimbriae minor subunit	5.2	6.9		
AAC76357-1	*gspK*	Putative type II secretion system protein K	5.2			
AAT48154-1	*yggP*	Uncharacterized protein	5.2	6.5	-4.9	
AAC75667-1	*ratB*	UPF0125 protein	5.2	6.3		
AAC74581-1	*hipB*	Antitoxin	5.1			
AAC75425-1	*dsdA*	D-serine dehydratase	5.1			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC74956-1	*tar*	Methyl-accepting chemotaxis protein II	5.1		-8.3	
AAC74307-1	*narK*	Nitrate/nitrite transporter	5.1	11.6	-5.4	
AAC75408-1	*intS*	Prophage integrase	5.1	15.9		
AAC75456-1	*yfeD*	Uncharacterized protein	5.1			
AAC76114-1	*ygjJ*	Uncharacterized protein	5.1		-4.3	
ABP93435-1	*ylcH*	Uncharacterized protein	5.1			
AAC75880-1	*araE*	Arabinose-proton symporter	5.0			
AAC74393-1	*ycjO*	Inner membrane ABC transporter permease protein	5.0		-4.4	
AAC74618-1	*pinQ*	Serine recombinase	5.0			
AAC75064-1	*yeeT*	Uncharacterized protein	5.0			
AAC76804-1	*aslA*	Arylsulfatase	4.9			
AAC76590-1	*xylF*	D-xylose-binding periplasmic protein	4.9			
EBE00001514921	*ycgH*	pseudogene	4.9	48.7		
EBE00001514887	*yedS*	pseudogene	4.9	4.3		
AAC76353-1	*gspG*	Putative type II secretion system protein G	4.9			
AAC77287-1	*kptA*	RNA 2'-phosphotransferase	4.9			
AAC75429-1	*evgS*	Sensor protein EvgS	4.9			
AAC74580-1	*hipA*	Serine/threonine-protein kinase toxin	4.9			
AAC75330-1	*yfbK*	Uncharacterized protein	4.9			
AAC73618-1	*allC*	Allantoate amidohydrolase	4.8			
AAC73658-1	*borD*	Lipoprotein Bor homolog from lambdoid prophage DLP12	4.8	40.0		
EBE00001514782	*ydbA*	pseudogene	4.8	7.0		
AAC74999-1	*yedL*	Uncharacterized N-acetyltransferase	4.8	9.6		
AAC73253-1	*folK*	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	4.7	19.7		
AAC73751-1	*hscC*	Chaperone protein	4.7	5.4		
AAC74961-1	*flhC*	Flagellar transcriptional regulator	4.7	6.0		

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC74962-1	*flhD*	Flagellar transcriptional regulator	4.7	15.1		
AAC73668-1	*ybcH*	Uncharacterized protein	4.7			
AAC76735-1	*yieE*	Uncharacterized protein	4.7			
AAC73399-1	*ykgM*	50S ribosomal protein L31 type B	4.6			
AAC74891-1	*mntP*	Probable manganese efflux pump	4.6	4.2		
AAC75365-1	*rpnB*	Recombination-promoting nuclease RpnB	4.6			
AAC75166-1	*rcnR*	Transcriptional repressor	4.6			
AAC74142-1	*yceO*	Uncharacterized protein	4.6			
AAC75455-1	*yfeC*	Uncharacterized protein	4.6	11.1		
AAC75881-1	*kduD*	2-dehydro-3-deoxy-D-gluconate 5-dehydrogenase	4.5	23.6		
AAC76687-1	*adeQ*	Adenine permease	4.5	8.1	-6.9	
AAC74992-1	*fliS*	Flagellar protein	4.5			
AAC76033-1	*hybO*	Hydrogenase-2 small chain	4.5			
AAC75823-1	*mazG*	Nucleoside triphosphate pyrophosphohydrolase	4.5			
AAC73783-1	*ybfP*	Uncharacterized lipoprotein	4.5			
AAC77254-1	*yjhH*	Uncharacterized lyase	4.5	4.2		
AAC74715-1	*ydhI*	Uncharacterized protein	4.5	11.3		
AAC75418-1	*yfdP*	Uncharacterized protein	4.5			
AAC75892-1	*ygeI*	Uncharacterized protein	4.5		-4.2	-30.4
AAC74990-1	*fliC*	Flagellin	4.4			
AAC73919-1	*gsiD*	Glutathione transport system permease protein	4.4			
AAC73563-1	*tomB*	Hha toxicity modulator TomB	4.4			
AAC73521-1	*pgpA*	Phosphatidylglycerophosphatase A	4.4			
AAC75387-1	*yfcA*	Probable membrane transporter protein	4.4			
AAC76879-1	*frvR*	Putative frv operon regulatory protein	4.4			
AAC76707-1	*yidP*	Uncharacterized HTH-type transcriptional regulator	4.4			
AAC77252-1	*yjhF*	Uncharacterized permease	4.4			
AAC73150-1	*caiA*	Crotonobetainyl-CoA reductase	4.3			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC75842-1	*fucA*	L-fuculose phosphate aldolase	4.3			
AAC74412-1	*ynaI*	Low conductance mechanosensitive channel	4.3	5.0		
AAC76280-1	*yhdE*	Maf-like protein	4.3	9.5		
AAT48174-1	*yhdX*	Putative amino-acid ABC transporter permease protein	4.3			
AAC75076-1	*yeeY*	Uncharacterized HTH-type transcriptional regulator	4.3			
AAT48205-1	*yieP*	Uncharacterized HTH-type transcriptional regulator	4.3			
AAC75436-1	*yfdY*	Uncharacterized protein	4.3	13.8		
AAC74379-1	*puuA*	Gamma-glutamylputrescine synthetase	4.2	5.2	-7.6	-7.4
EBE00001514944	*ykiA*	pseudogene	4.2	6.6		
AAC77276-1	*fimH*	Type 1 fimbrin D-mannose specific adhesin	4.2			
AAC74932-1	*yebB*	Uncharacterized protein	4.2			
AAC75210-1	*mglA*	Galactose/methyl galactoside import ATP-binding protein	4.1			
EBE00001514967	*yjhD*	pseudogene	4.1			
EBE00001514960	*yfcU*	pseudogene	4.1			
AAC76354-1	*gspH*	Putative type II secretion system protein H	4.1		-4.3	
AAC74027-1	*elfG*	Uncharacterized fimbrial-like protein	4.1			
ABP93437-1	*ybfK*	Uncharacterized protein	4.1			
AAC77253-1	*yjhG*	Uncharacterized protein	4.1	5.7		
AAC75068-1	*yeeX*	UPF0265 protein	4.1			
AAC76107-1	*aer*	Aerotaxis receptor	4.0			
AAC77271-1	*fimI*	Fimbrin-like protein	4.0			
AAC73329-1	*yafQ*	mRNA interferase toxin	4.0	5.9		
AAC73304-1	*yaeF*	Probable lipoprotein peptidase	4.0			
AAC74866-1	*yoaG*	DUF1869 domain-containing protein	4.0			
AAC75878-1	*lysR*	Transcriptional activator protein	4.0			

Transcript id	Gene	Protein/Function	Fold-change (48h/24h) in WT		Fold-change (48h/24h) in Δ*rpoS*	
			RNA -seq	Micro array	RNA -seq	Micro array
AAC76639-1	*yibD*	Uncharacterized glycosyltransferase	4.0			
AAC75180-1	*yehL*	Uncharacterized protein	4.0			

(The transcripts also present in microarray are mention with fold-change value. The transcripts also showed higher abundance in Δ*rpoS* were RpoS-independent and fold-change are mention. The transcripts showed low abundance in Δ*rpoS* were RpoS-dependent and fold-change are mention with negative (-) sign. All the transcripts were with fold-change ≥ 4 and FDR adjusted *p*-value ≤ 0.05).

APPENDIX 7: Real Time PCR validation experiments

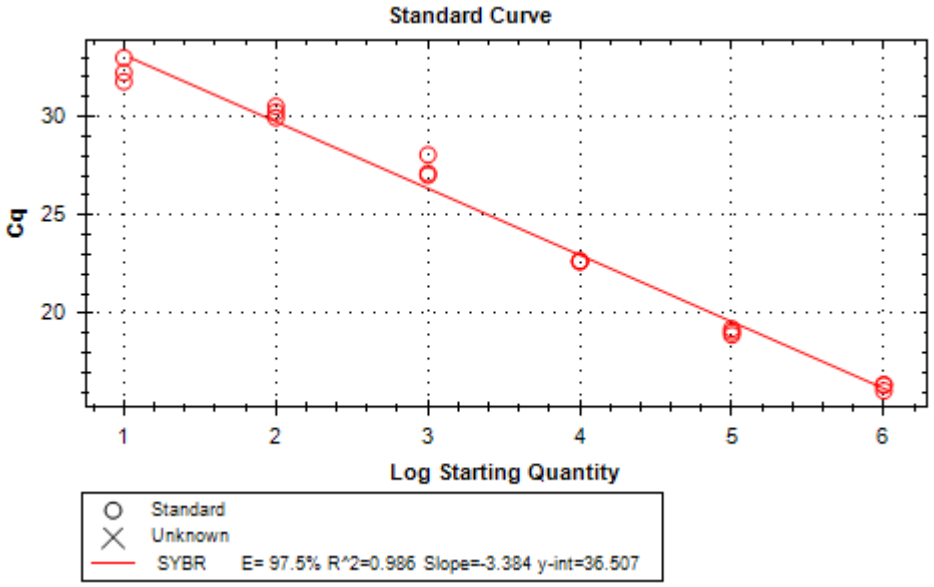


Figure 36: qPCR standard curve to test for amplification efficiency of the *rrsA* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *rrsA* gene amplicon.

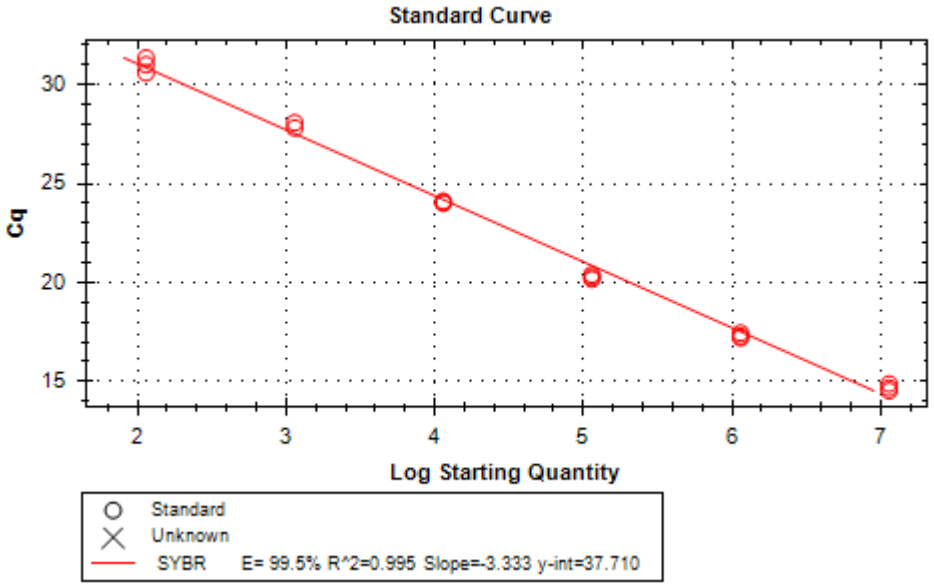


Figure 37: qPCR standard curve to test for amplification efficiency of the *entC* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *entC* gene amplicon.

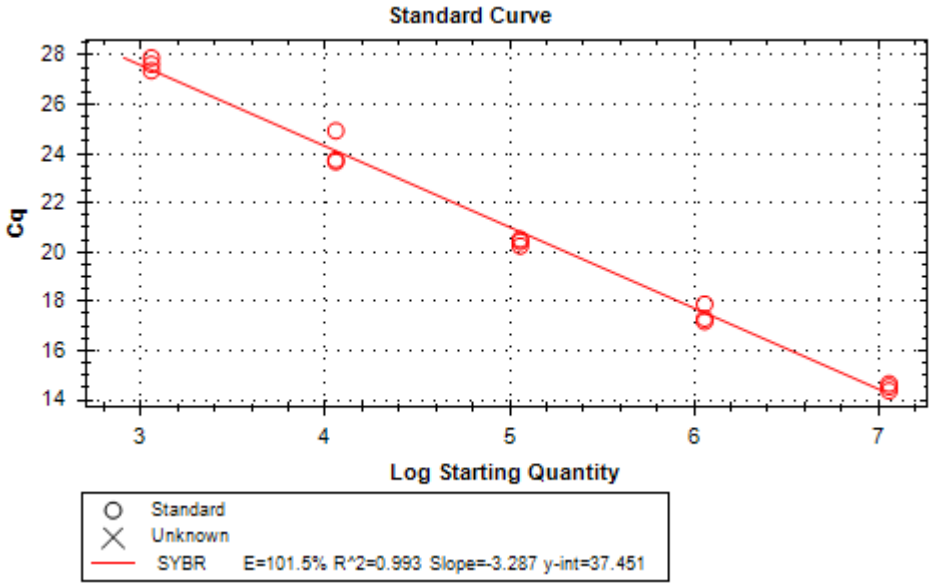


Figure 38: qPCR standard curve to test for amplification efficiency of the *fecR* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *fecR* gene amplicon.

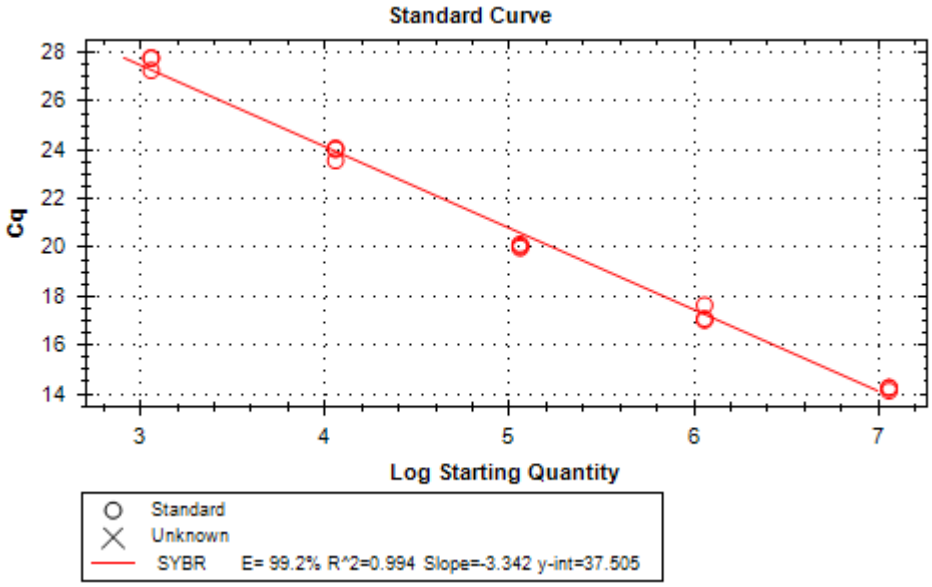


Figure 39: qPCR standard curve to test for amplification efficiency of the *entF* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *entF* gene amplicon.

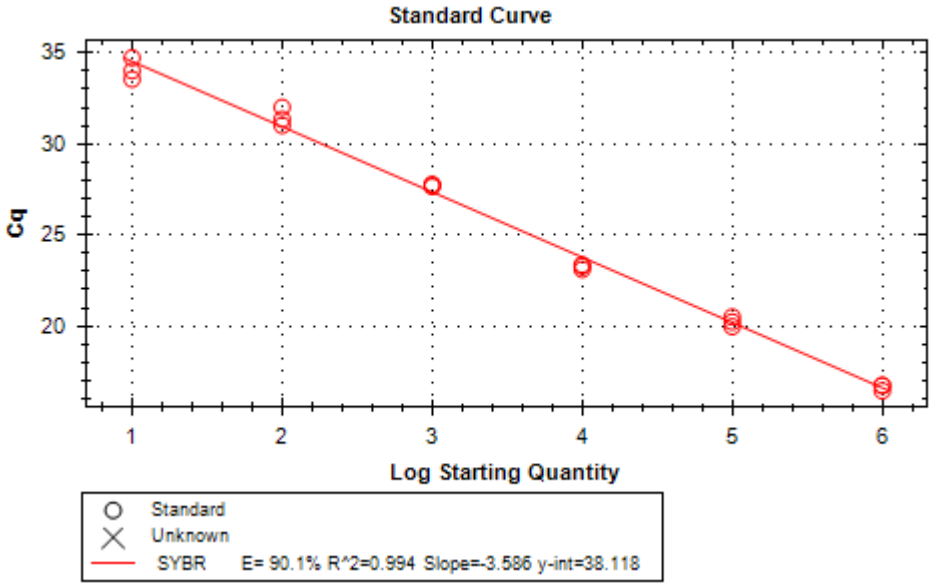


Figure 40: qPCR standard curve to test for amplification efficiency of the *fecI* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *fecI* gene amplicon.

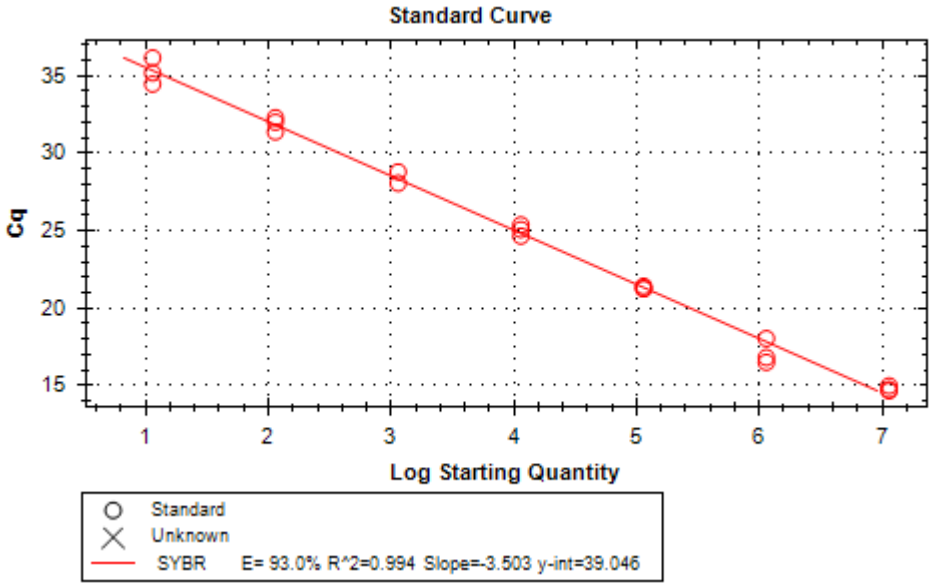


Figure 41: qPCR standard curve to test for amplification efficiency of the *astA* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *astA* gene amplicon.

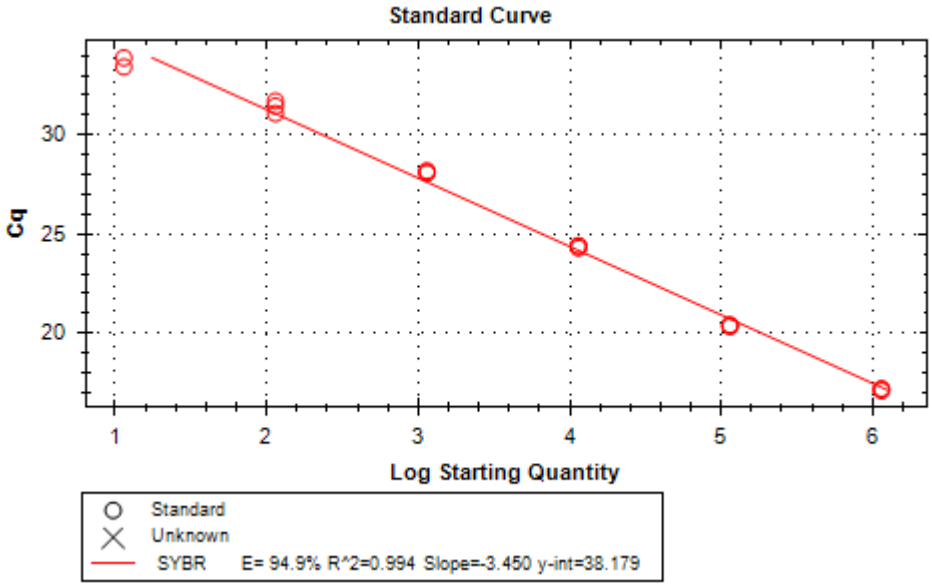


Figure 42: qPCR standard curve to test for amplification efficiency of the *astC* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *astC* gene amplicon.

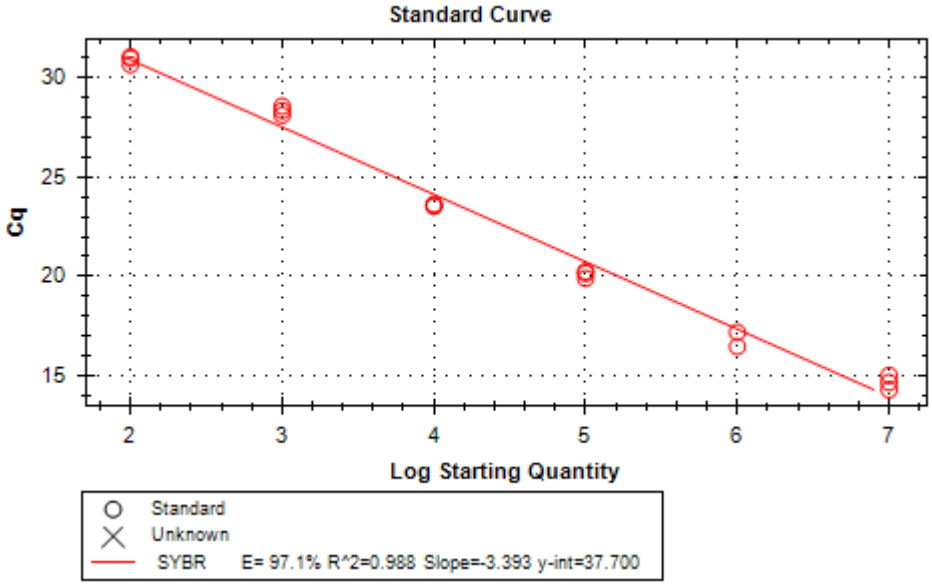


Figure 43: qPCR standard curve to test for amplification efficiency of the *yadN* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *yadN* gene amplicon.

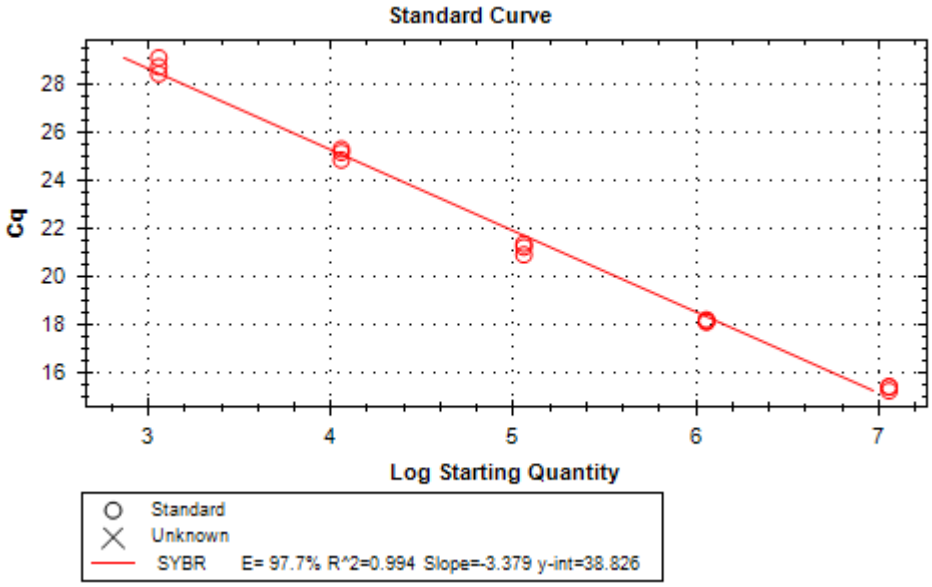


Figure 44: qPCR standard curve to test for amplification efficiency of the *yadV* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *yadV* gene amplicon.

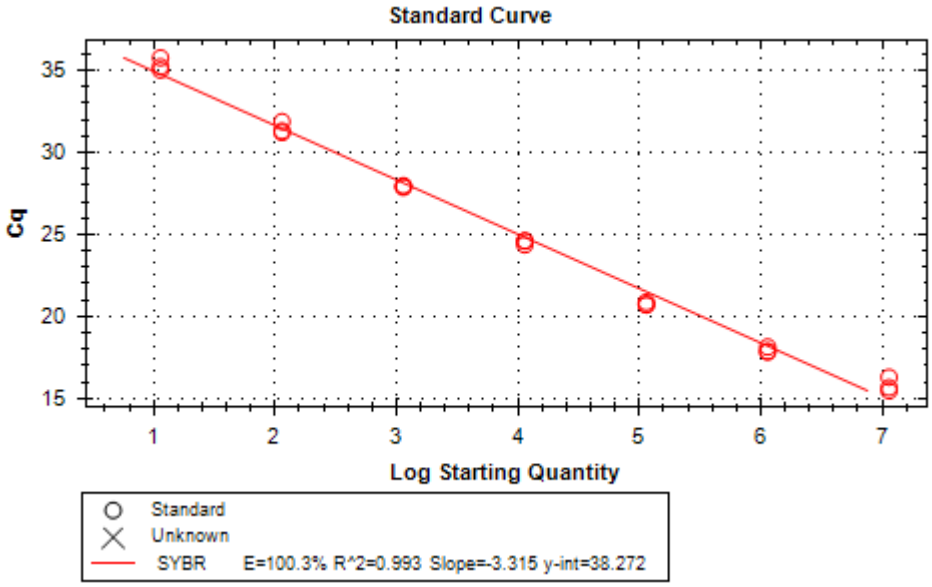


Figure 45: qPCR standard curve to test for amplification efficiency of the *sfmH* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *sfmH* gene amplicon.

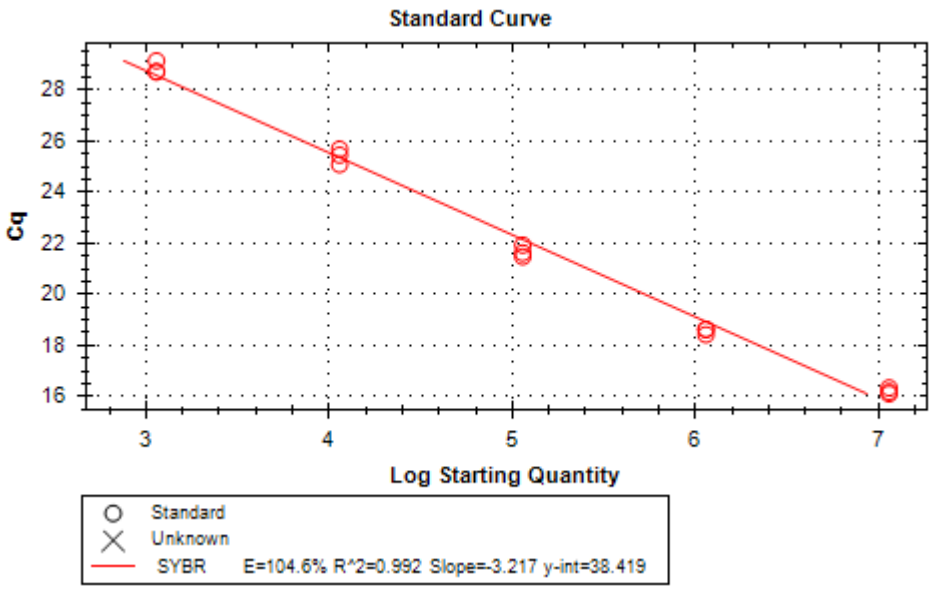


Figure 46: qPCR standard curve to test for amplification efficiency of the *mqsR* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *mqsR* gene amplicon.

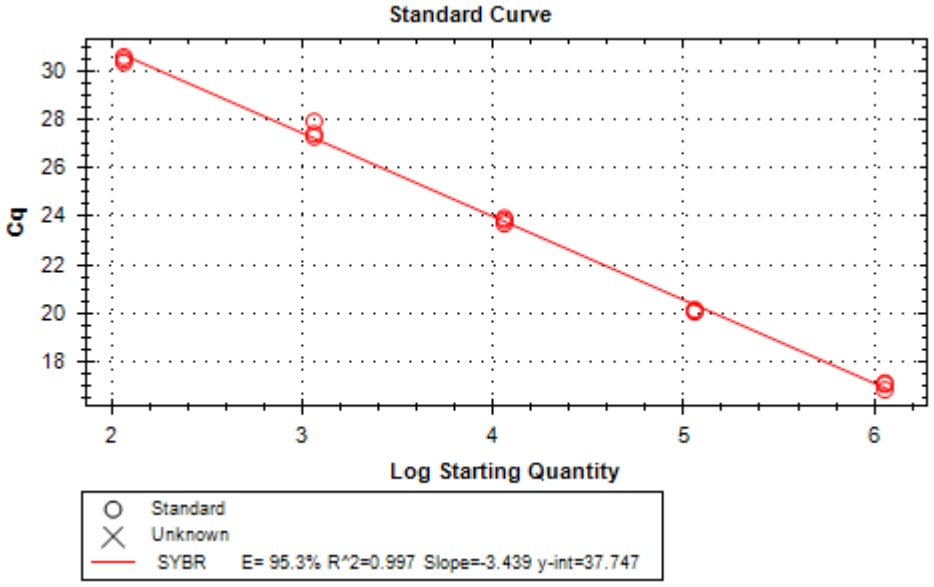


Figure 47: qPCR standard curve to test for amplification efficiency of the *mqsA* amplicon. Standard curve using 10-fold serial dilution of K12 MG1655 genomic DNA to test for efficiency of amplification of the *mqsA*gene amplicon.

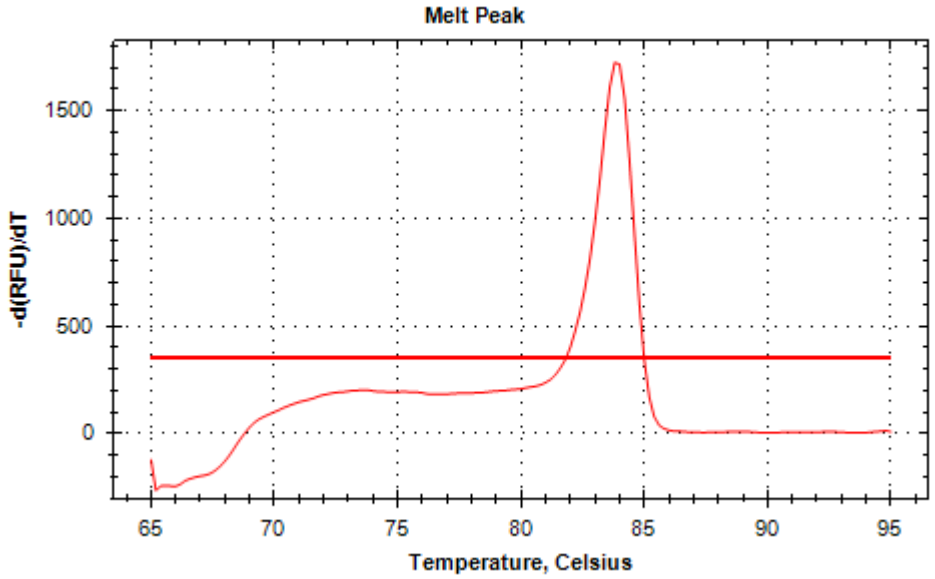


Figure 48: Melt curve analysis of the rrsA amplicon.
Single melt curve peak indicates that the amplification is specific.

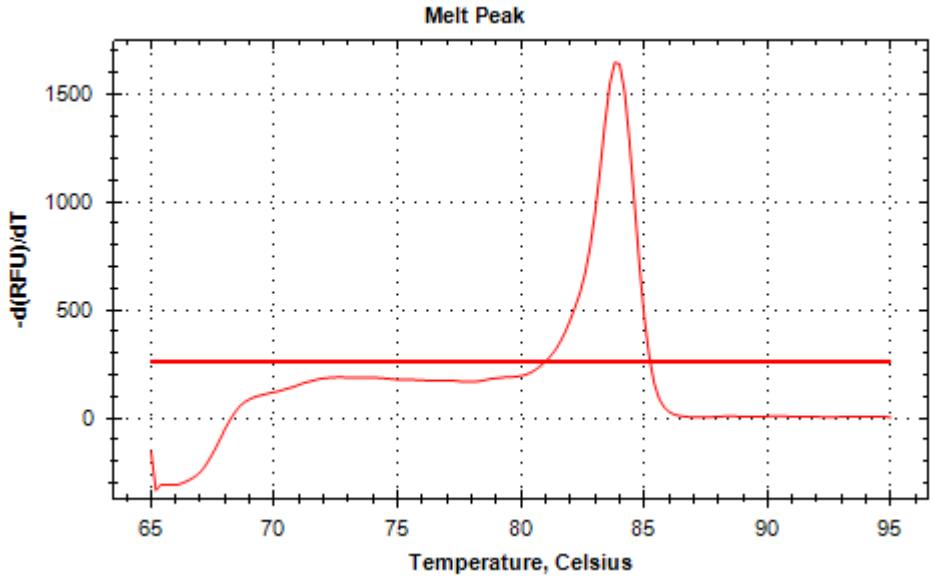


Figure 49: Melt curve analysis of the *entC* amplicon.
Single melt curve peak indicates that the amplification is specific.

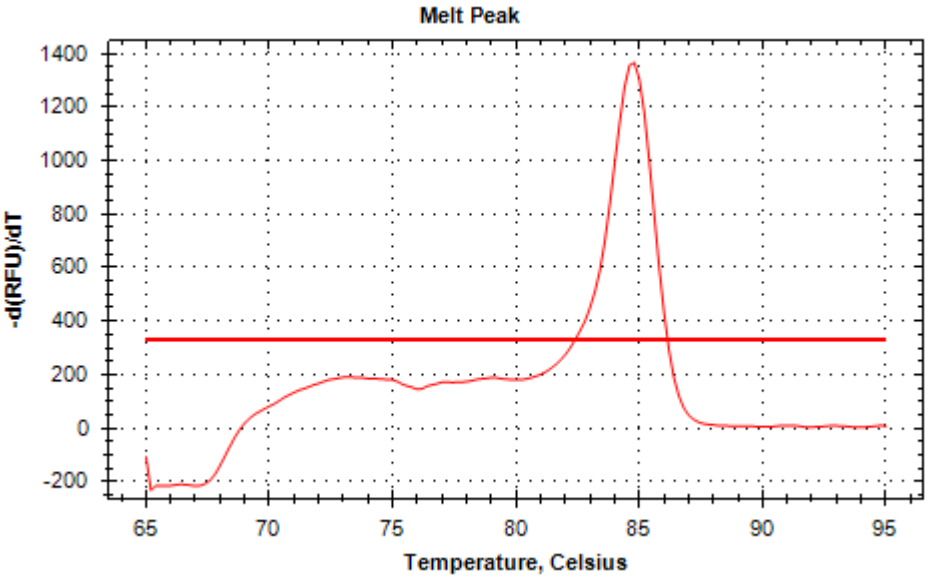


Figure 50: Melt curve analysis of the *fecR* amplicon.
Single melt curve peak indicates that the amplification is specific.

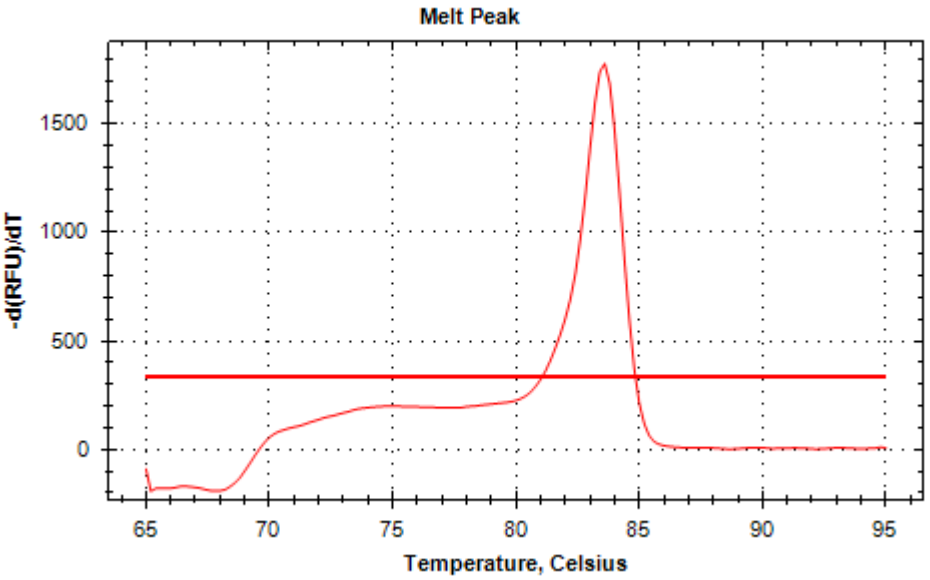


Figure 51: Melt curve analysis of the *entF* amplicon.
Single melt curve peak indicates that the amplification is specific.

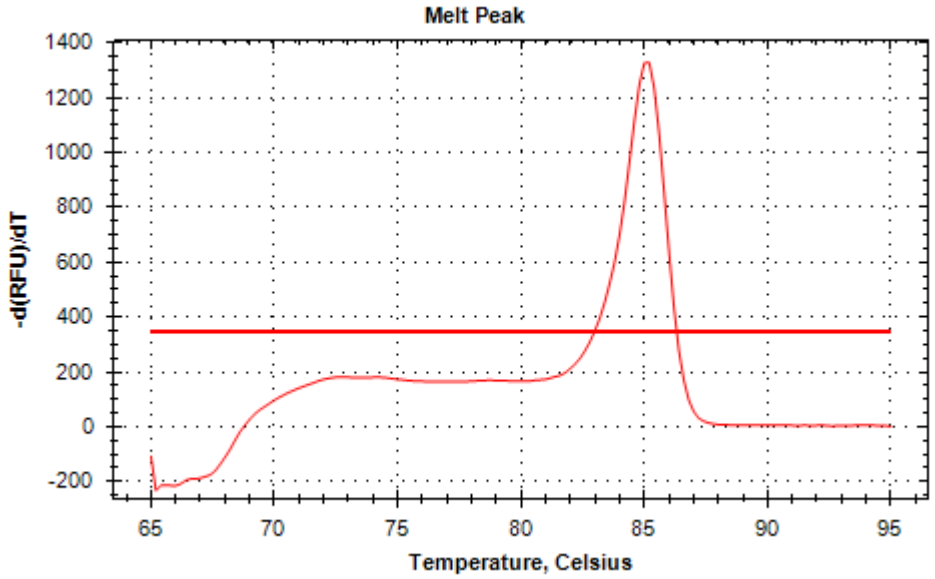


Figure 52: Melt curve analysis of the *fecI* amplicon.
Single melt curve peak indicates that the amplification is specific.

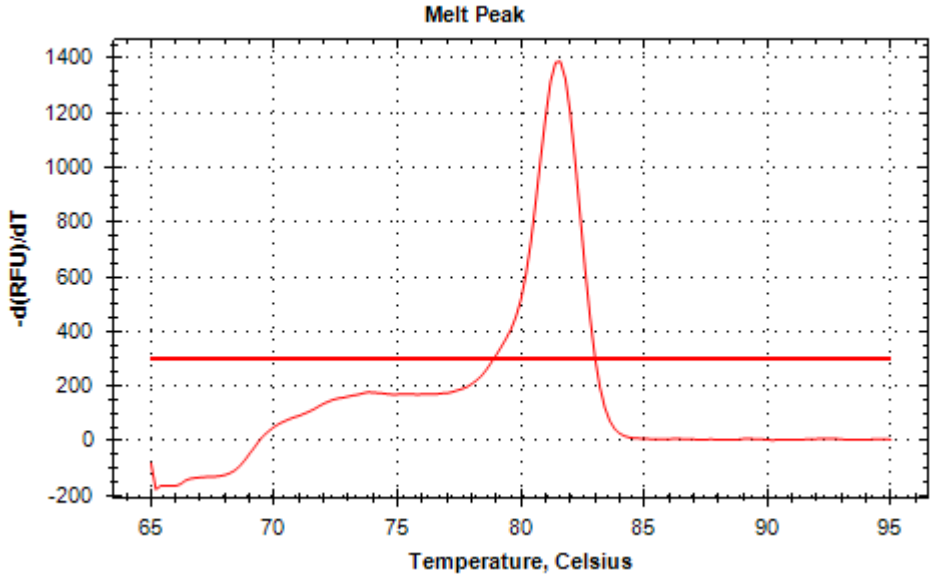


Figure 53: Melt curve analysis of the *astA* amplicon.
Single melt curve peak indicates that the amplification is specific.

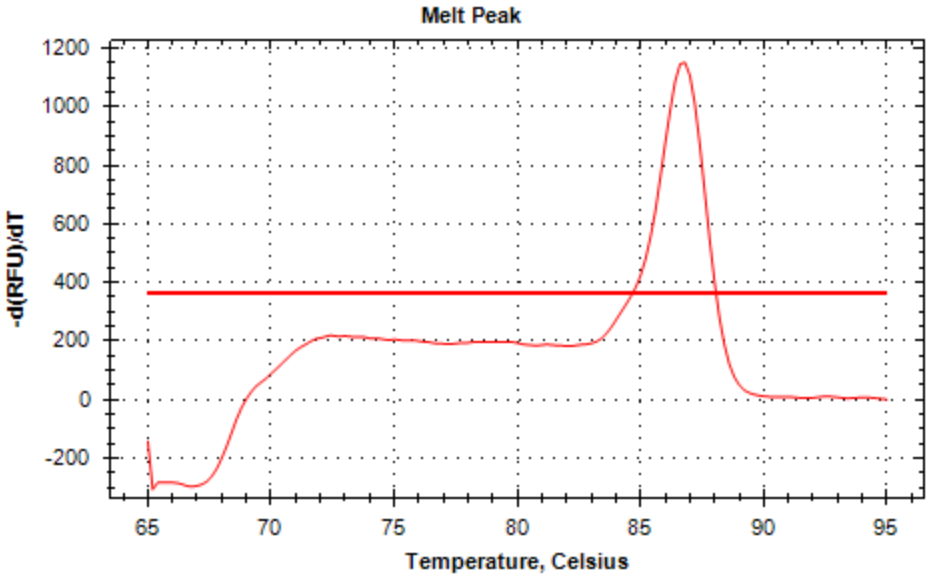


Figure 54: Melt curve analysis of the *astC* amplicon.
Single melt curve peak indicates that the amplification is specific.

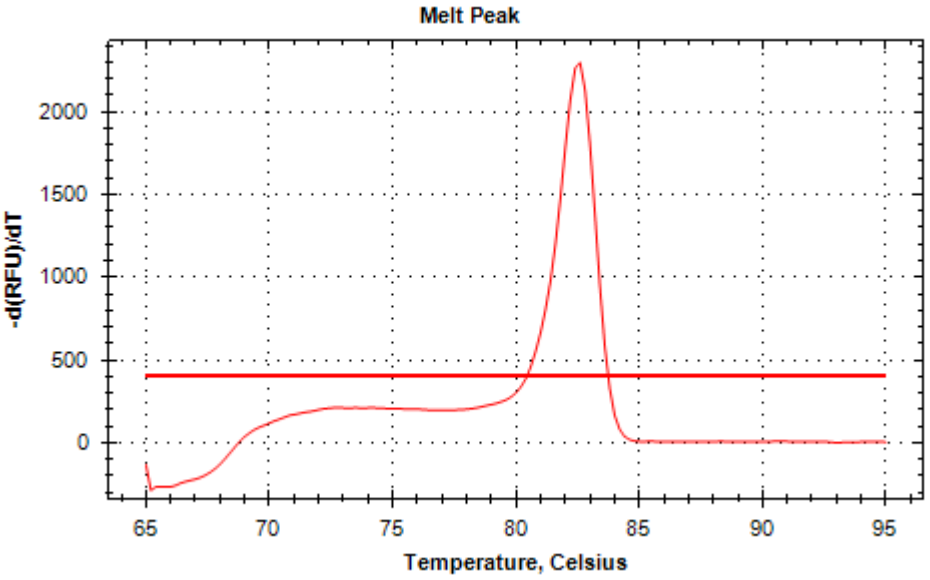


Figure 55: Melt curve analysis of the *yadN* amplicon.
Single melt curve peak indicates that the amplification is specific.

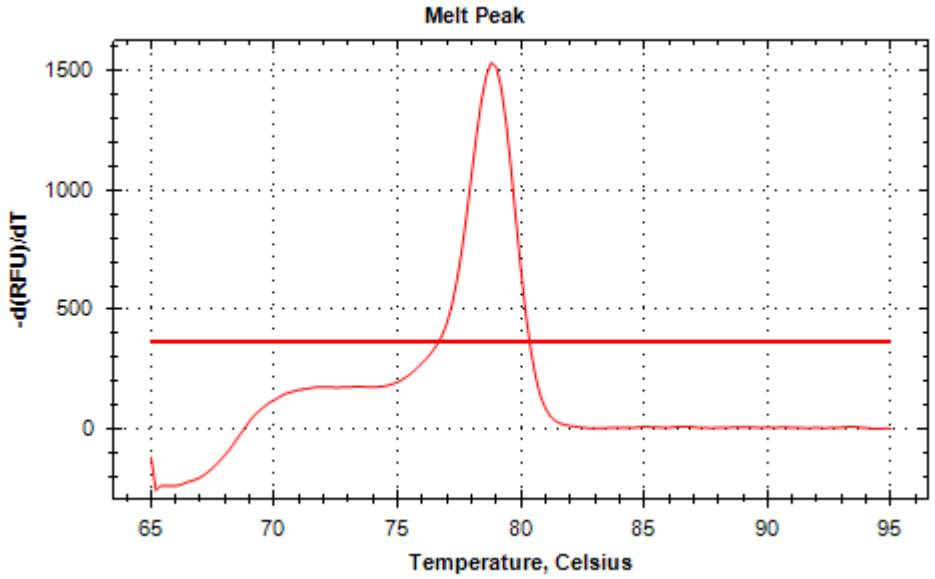


Figure 56: Melt curve analysis of the *yadV* amplicon.
Single melt curve peak indicates that the amplification is specific.

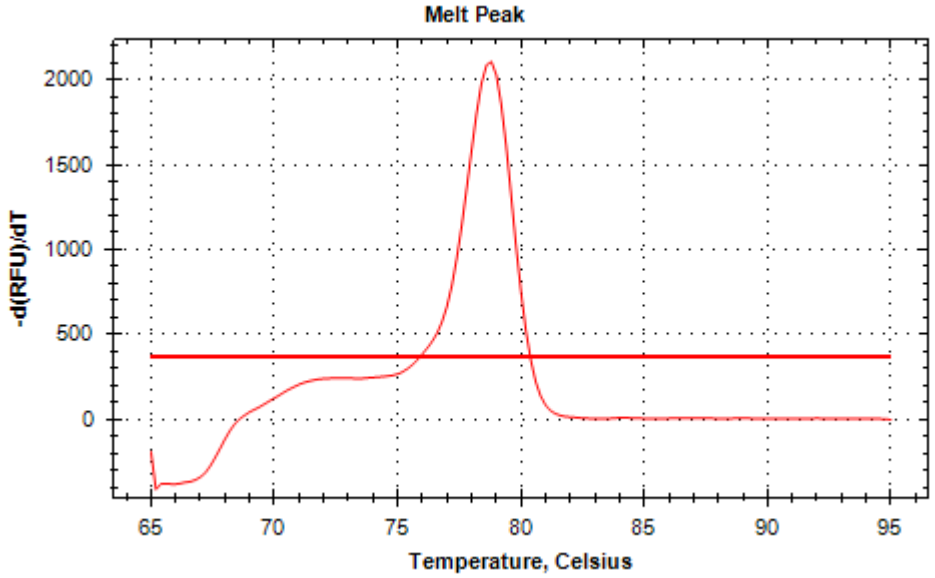


Figure 57: Melt curve analysis of the *sfmH* amplicon.
Single melt curve peak indicates that the amplification is specific.

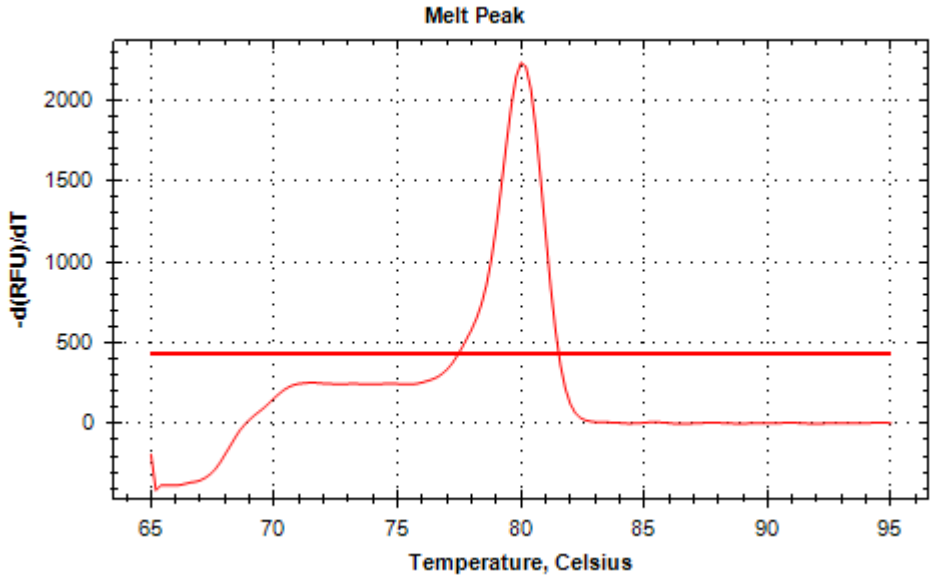


Figure 58: Melt curve analysis of the *mqsR* amplicon.
Single melt curve peak indicates that the amplification is specific.

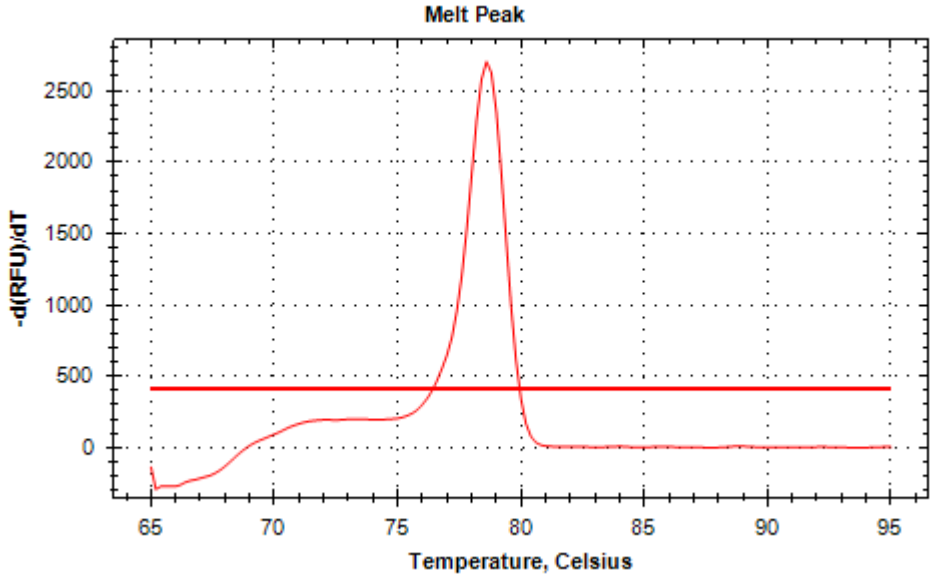


Figure 59: Melt curve analysis of the *mqsA* amplicon.
Single melt curve peak indicates that the amplification is specific.

APPENDIX 8: Raw data for qPCR

Table 15: Raw Cq values corresponding to data presented in validation data for gene expression analysis in WT (*E. coli* K12 MG1655).

Growth phase	Gene name	Rep 1	Rep 2	Rep 3
0.3OD	*fecR*	32.63	32.36	32.17
1.5OD	*fecR*	30.11	30.19	30.04
24h	*fecR*	27.46	27.55	27.42
48h	*fecR*	30.88	30.89	30.62
0.3OD	*fecI*	31.66	31.86	31.57
1.5OD	*fecI*	29.19	29.09	29.37
24h	*fecI*	26.65	26.68	26.88
48h	*fecI*	30.51	31.54	30.89
0.3OD	*astA*	32.88	32.73	32.53
1.5OD	*astA*	27.88	27.89	27.72
24h	*astA*	25.04	24.78	27.28
48h	*astA*	28.31	28.04	27.85
0.3OD	*astC*	33.73	34.18	34.15
1.5OD	*astC*	27.08	27.01	27.13
24h	*astC*	24.72	24.80	24.88
48h	*astC*	26.62	26.54	26.56
0.3OD	*entC*	32.59	32.65	32.10
1.5OD	*entC*	28.96	28.95	29.12
24h	*entC*	27.28	27.39	27.48
48h	*entC*	29.12	29.07	29.11
0.3OD	*entF*	32.10	32.08	31.91
1.5OD	*entF*	30.88	31.15	30.75
24h	*entF*	28.15	28.22	28.00
48h	*entF*	30.03	30.24	29.99
0.3OD	*yadN*	33.98	33.75	33.83
1.5OD	*yadN*	34.68	34.09	34.01
24h	*yadN*	35.23	35.27	35.31
48h	*yadN*	28.73	28.84	28.83
0.3OD	*yadV*	34.33	35.11	35.16
1.5OD	*yadV*	36.41	37.29	36.38
24h	*yadV*	37.40	36.94	36.63
48h	*yadV*	30.02	30.19	30.13
0.3OD	*sfmH*	36.13	36.58	35.95
1.5OD	*sfmH*	36.42	35.71	36.23
24h	*sfmH*	35.78	36.37	36.23
48h	*sfmH*	30.69	31.03	30.93
0.3OD	*mqsR*	29.10	28.79	28.64
1.5OD	*mqsR*	27.18	27.25	27.06
24h	*mqsR*	25.41	25.61	25.34
48h	*mqsR*	24.69	24.60	24.49
0.3OD	*mqsA*	30.69	30.66	31.07

1.5OD	*mqsA*	29.39	29.00	29.16
24h	*mqsA*	27.02	27.11	27.45
48h	*mqsA*	25.05	24.98	25.96

Exponential phase - 0.3 OD, Early stationary - 1.5 OD and Prolonged-incubation phase -24h and 48h

Table 16: Raw Cq values corresponding to data presented in validation graphs for gene expression analysis in Δ*rpoS* (isogenic mutant of *E. coli* K12 MG1655).

Growth phase	Gene name	Rep 1	Rep 2	Rep 3
0.3OD	*fecR*	29.05	29.12	29.10
1.5OD	*fecR*	31.84	31.74	32.08
24h	*fecR*	25.29	25.35	25.38
48h	*fecR*	30.50	30.46	30.72
0.3OD	*fecI*	29.28	29.02	28.86
1.5OD	*fecI*	31.58	31.49	31.25
24h	*fecI*	25.20	25.31	25.15
48h	*fecI*	29.43	29.46	29.37
0.3OD	*astA*	33.05	33.12	33.15
1.5OD	*astA*	30.02	30.27	29.82
24h	*astA*	24.98	25.14	25.00
48h	*astA*	29.12	29.01	29.03
0.3OD	*astC*	33.31	34.03	33.21
1.5OD	*astC*	29.54	29.47	29.54
24h	*astC*	24.46	24.52	24.77
48h	*astC*	29.15	29.16	29.12
0.3OD	*entC*	26.63	26.63	26.74
1.5OD	*entC*	32.81	32.58	32.71
24h	*entC*	24.81	24.80	25.09
48h	*entC*	27.37	27.36	27.50
0.3OD	*entF*	27.69	25.35	25.11
1.5OD	*entF*	31.35	31.63	32.24
24h	*entF*	24.61	24.59	24.67
48h	*entF*	27.66	27.63	27.87
0.3OD	*fimB*	27.61	27.53	28.02
1.5OD	*fimB*	29.05	29.25	29.30
24h	*fimB*	24.62	24.57	24.69
48h	*fimB*	30.05	30.15	30.26
0.3OD	*yadN*	35.49	35.02	34.83
1.5OD	*yadN*	36.50	35.96	36.24
24h	*yadN*	25.87	25.97	25.74
48h	*yadN*	34.81	34.27	33.81
0.3OD	*yadV*	34.26	33.92	34.21
1.5OD	*yadV*	35.56	35.12	35.01
24h	*yadV*	24.81	25.02	25.17
48h	*yadV*	34.89	33.54	34.22
0.3OD	*yehC*	31.06	31.41	31.31

1.5OD	*yehC*	32.59	32.67	32.49
24h	*yehC*	25.88	25.84	26.43
48h	*yehC*	34.63	34.19	35.19
0.3OD	*sfmH*	35.91	35.87	35.57
1.5OD	*sfmH*	34.36	35.11	35.01
24h	*sfmH*	25.20	25.41	25.51
48h	*sfmH*	35.20	34.46	34.45
0.3OD	*mqsR*	27.62	27.29	27.26
1.5OD	*mqsR*	24.97	25.14	24.92
24h	*mqsR*	24.43	24.59	24.37
48h	*mqsR*	24.43	24.48	24.39
0.3OD	*mqsA*	29.49	29.73	29.51
1.5OD	*mqsA*	26.94	26.80	26.80
24h	*mqsA*	24.50	24.56	24.75
48h	*mqsA*	26.50	26.55	26.67

Exponential phase - 0.3 OD, Early stationary - 1.5 OD and Prolonged-incubation phase -24h and 48h

APPENDIX 9: Primer sequences used for RT-qPCR study

Table 17: Sequences of primers used in this study.

Primers	Sequence (5' – 3')
fecI_F	GCGCTGGAAAAAGCGTATC
fecI_R	CATGCTGTCGAGGAGTTGTAG
fecR_F	TTACTACCGCGAAAGATGCC
fecR_R	CTGGCGGACGGTAAATTCTG
entC_F	GACTCAGGCGATGAAAGAGG
entC_R	TCAAAGGGAGTTGCGAGATG
entF_F	CTTCGTGAAACATTGCCACC
entF_R	TCAGTTCAGGCAACGGTAAG
astA_F	CCTGGTACAACTATCGCGTC
astA_R	TTACTGAGAAACAGCGTCGG
astC_F	GTATATCGACTTCGCGGGTG
astC_R	ACTTACTCGCCTGTTCGTTC
bssR_F	CGTCAGCGAAAGCAATCATC
bssR_R	AGAGCACTCCACTCTTCCTG
fimB_F	AATCCGCTTTCTCGGCAAC
fimB_R	ATTCGCCAAAGCAAAACCAC
yadN_F	ATGCTGGCGTACTGAATGAC
yadN_R	CATGTCGTTGTTCAAAGTCCC
yadV_F	CCAAACGTGGGCAAACAATC
yadV_R	CAGAACACGCTCTCTCTGTC
sfmH_F	CGATGGCATGTTTGTGTCTG
sfmH_R	TGTTCATCTTCTGCTACGCC
mqsR_F	AAAACGCACACCACATACAC
mqsR_R	CTGCATTTAACAGGGCACTAC
mqsA_F	TTTGCCACCAGGGAGAAATG
mqsA_R	ATGCTCTCTTCGCAATGGAC
rrsA_F	AGATGAGAATGTGCCTTCGG
rrsA_R	CGCTGGCAACAAAGGATAAG

F and R in primer name indicates forward and reverse primer respectively

APPENDIX 10: RpoS regulon member peak during the early stationary phase and decline during the prolonged-incubation phase

Table 18: RpoS regulon member peak during the early stationary phase and decline during the prolonged-incubation phase in WT.

Transcript id	Protein/Function	Gene	Fold-change (1.5/0.3)		Fold-change (48h/24h)	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC74566-1	Glutamate decarboxylase beta	*gadB*	90.43	103.90	-22.27	-37.43
AAC74565-1	Probable glutamate/gamma-aminobutyrate antiporter	*gadC*	87.20	91.71	-16.20	-37.05
AAC76542-1	Glutamate decarboxylase alpha	*gadA*	83.87	95.04	-15.88	-45.31
AAC73588-1	Inner membrane transport protein	*ybaT*	37.95	42.08	-10.22	**-2.39**
AAC73840-1	Uncharacterized protein	*ybgS*	35.50	31.95	**-3.64**	**-1.52**
AAC73587-1	Glutaminase 1	*glsA*	34.90	33.96	-19.28	-9.67
AAC76537-1	Transcriptional regulator	*gadE*	30.54	14.08	-27.55	-32.67
AAC74341-1	Uncharacterized protein	*yciG*	28.98	4.63	**-3.17**	
AAC75033-1	Protein deglycase 1	*hchA*	27.26	22.59	-6.02	-9.37
AAC76533-1	Putative magnesium transporter	*yhiD*	24.66	43.20	**-3.74**	
AAC76579-1	Uncharacterized HTH-type transcriptional regulator	*yiaG*	23.83	22.38	**-2.63**	
AAC76534-1	Acid stress chaperone	*hdeB*	23.59	**2.43**	-34.63	-8.45
AAT48137-1	Catalase HPII	*katE*	23.17	19.02	-6.46	-5.74
AAC76535-1	Acid stress chaperone	*hdeA*	20.72	**2.13**	-29.57	-9.04
AAC76536-1	Acid resistance membrane protein	*hdeD*	20.27	38.59	-6.93	-6.79
AAC75710-1	Gamma-aminobutyrate permease	*gabP*	19.99	16.63	-4.06	-50.86
AAC77329-1	Osmotically-inducible protein Y	*osmY*	19.80	23.52	-4.90	-3.73
AAC76060-1	BOF family protein	*ygiW*	19.61	20.63	**-3.62**	
ABD18720-1	UPF0391 membrane protein	*ytjA*	16.72	14.83	-6.39	
AAC74088-1	Uncharacterized protein	*yccJ*	15.54	26.39	**-2.31**	**-1.33**
AAC74057-1	Hydrogenase-1 small chain (NiFe hydrogenase)	*hyaA*	15.03	11.51	**-2.25**	-6.49
AAC77015-1	UPF0337 protein	*yjbJ*	14.00	9.14	-4.96	

Transcript id	Protein/Function	Gene	Fold-change (1.5/0.3)		Fold-change (48h/24h)	
			RNA-seq	Micro array	RNA-seq	Micro array
AAC76039-1	Uncharacterized oxidoreductase	*yghA*	13.96	**2.28**	-3.33	
AAC74966-1	Trehalose-6-phosphate synthase	*otsA*	13.45	15.49	-1.14	**-2.41**
AAC74134-1	Uncharacterized protein	*yceK*	13.20	11.47	-2.68	
AAC74058-1	Hydrogenase-1 large chain (NiFe hydrogenase)	*hyaB*	12.75	15.09	-3.25	-6.24
AAC74555-1	Osmotically-inducible protein C	*osmC*	12.47	9.91	-2.76	-4.10
AAC73899-1	DNA protection during starvation protein	*dps*	12.33	8.82	-5.09	-61.98
AAC74135-1	Acidic protein	*msyB*	11.45	13.29	-2.27	-4.72
AAC73432-1	Uncharacterized protein	*yahO*	9.45	10.73	-2.76	**-2.40**
AAC74549-1	Respiratory nitrate reductase 2 beta chain	*narY*	9.12	4.02	-3.14	-4.28
AAC73324-1	Inhibitor of vertebrate lysozyme	*ivy*	7.49	7.31	-2.27	-6.02
AAC74550-1	Respiratory nitrate reductase 2 alpha chain	*narZ*	7.01	**2.13**	-2.63	-4.13
AAT48242-1	Entericidin B	*ecnB*	4.29	14.15	-1.87	-56.83

(The transcripts also present in microarray are mention with fold-change value. The transcripts showed low abundance in during prolonged-incubation are mention with negative (-) sign. All the transcripts are significant with FDR adjusted p-value ≤ 0.05. Bold numbers are less than 4-fold change).

APPENDIX 11: Correlation graph of RNA-seq data in WT

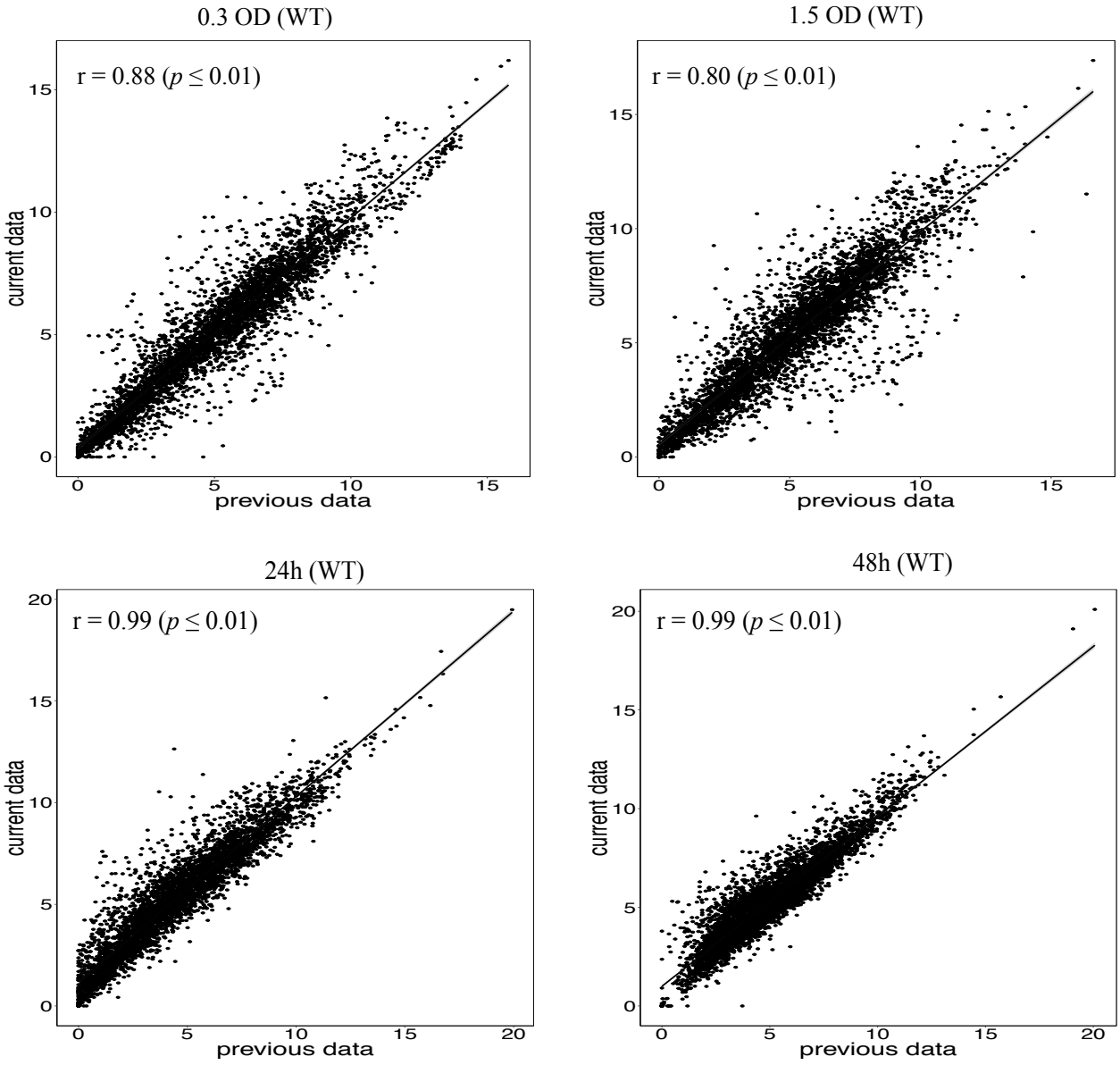


Figure 60: Pearson correlation between current and previous RNA-sequencing data in WT.
The relationship between the expression profiles generated by RNA-seq is depicted as a linear regression line. Pearson correlation coefficient represented by r value, and *p*-value shows the level of significance. Previous data was generated in our lab by previous student using same RNA extraction and analysis method.

Standard Operating Procedures

Bacterial growth

1) Streak strains from the -80 °C glycerol stock cultures without thawing onto LB plates.
2) Inoculate a single colony into 10 ml of LB in a 50 ml Erlenmeyer flask and incubate at 37 °C aerobically with shaking at 200 rpm (Innova 4000, New Brunswick Scientific).
3) After overnight growth (typically 12 h) subculture 1:10, 000 into prewarmed 50-ml LB in 250-ml flasks and monitor OD600 using Multiskan Spectrum (Thermo Labsystems).
4) Exponential phase is defined as OD600 = 0.3 (typically 4 h post inoculation), early stationary phase as OD600 = 1.5 (5 h 10 min post inoculation) and prolonged incubation as 24h and 48h old culture.

Total RNA extraction (using Norgen Biotek, Cat No: 37500)

1) Grow the cells to appropriate density/time and pellet approx. 10^9 cells at 14, 000 g for 2 min.
2) Discard the supernatant with a pipette and re-suspend the pellet in 100 µl of 1 mg/ml lysozyme in TE buffer. Incubate for 5 min.
3) Add 300 µl of Buffer RL and vortex to mix thoroughly.
4) Add 200 µl of anhydrous ethanol and vortex to mix thoroughly.
5) Transfer the 600 µl of the mixture to the RNA extraction column and centrifuge at 10, 000 g for 1 min.
6) Add 400 µl of wash solution to the column and centrifuge for 1 min.
7) Wash the column twice (a total of three times) with 400 µl of wash solution.
8) Centrifuge for 2 min at 14, 000 g to dry the column.
9) Add 50 µl of the elution solution directly on top of the column and centrifuge at 200 g for 2 minutes.
10) Centrifuge at 14, 000 g for 1 min to collect the RNA. RNA can be quantified using Invitrogen Qubit reagent for RNA quantification (Q32855) or Nanodrop 2000 (the former is considered a more accurate method of quantification).

In-solution DNase 1 treatment of RNA

1) For every 1 µg of RNA 1 U of DNase 1 is used for digestion at 37 °C for 30 min. Combine the reagents as shown in the following table. Scale up the reaction if more RNA is to be digested:

Reagent	Volume/Amount
RNA	1.0 µg
RNase free DNase 1 (EN0521)	1 U (1.0 µl)
10 X DNase 1 Buffer	1.0 µl
RNase free H_2O	Up to 10.0 µl

Repurification of RNA sample following DNase 1 treatment (using RNA Clean and Concentrator – Zymo Research Cat. No. R1015)

1) Add 2 volumes of the RNA binding buffer to the DNase 1 treatment reaction mixture and mix well.
2) Add an equal volume of anhydrous ethanol to the reaction mixture and mix well.
3) Add the above mixture to the **Zymo Spin** column and centrifuge at 10, 000 g for 30 s.
4) Add 400 µl of RNA prep buffer to the column and centrifuge at 10, 000 g for 30 s.
5) Add 700 µl of RNA wash buffer to the column and centrifuge at 10, 000 g for 2 min to dry the column.
6) Add the desired amount of RNase-free water to the column and centrifuge at 10, 000 g for 30 s. Measure RNA concentration.

cDNA synthesis

1) Combine the following reagents in a 0.2 ml PCR tube a

Reagent	Volume/Amount
RNA	500.0 ng
5X Iscript cDNA synthesis mix (*BIORAD*, 170-8890)	4.0 µl
RNase-free water	To 20.0 µl

2) Incubate the tube in a thermal cycler under the following conditions: 1) 5 min at 25 °C 2) 30 min at 42 °C 3) 5 min at 85 °C 4) hold at 4 °C/store at 4 °C.

Quantitative PCR (qPCR)

The following are the reaction volumes for a 10-µl reaction (to prepare a master mix, multiply each volume/amount in the following table by the number of reactions to be conducted).

Reagent	Volume/Amount
2X Sso-Fast reaction mix (*BIORAD*, 172-5200)	5.0 µl
Forward primer (25 µM)	0.2 µl
Reverse primer (25 µM)	0.2 µl

cDNA template	1.0 µl
ddH_2O (PCR grade)	To 10.0 µl

Thermal cycling protocol for cDNA quantification: 1) 95 °C for 2 minutes (initial denaturation) 2) 95 °C for 0.05 s 3) annealing and extension at 55 °C for 10 s. Repeat steps 2), 3), and 4) for a total of 40 times. Record fluorescence after each cycle after step 3). Increase the temperature at 0.05 °C intervals from 65 °C to 95 °C recording fluorescence at each interval increase to generate the melt curve.

References

Agladze, K., X. Wang & T. Romeo, (2005) Spatial periodicity of *Escherichia coli* K-12 biofilm microstructure initiates during a reversible, polar attachment phase of development and requires the polysaccharide adhesin PGA. *Journal of Bacteriology* **187**: 8237-8246.

Akesson, M., E.N. Karlsson, P. Hagander, J. Axelsson & A. Tocaj, (1999) On-line detection of acetate formation in *Escherichia coli* cultures using dissolved oxygen responses to feed transients. *Biotechnology and Bioengineering* **64**: 590-598.

Amores, G.R., A. de las Heras, A. Sanches-Medeiros, A. Elfick & R. Silva-Rocha, (2017) Systematic identification of novel regulatory interactions controlling biofilm formation in the bacterium *Escherichia coli*. *Scientific Reports* **7**: 16768.

Anders, S. & W. Huber, (2010) Differential expression analysis for sequence count data. *Genome Biology* **11**: R106.

Anders, S., P.T. Pyl & W. Huber, (2015) HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166-169.

Andersen, J.B., C. Sternberg, L.K. Poulsen, S.P. Bjørn, M. Givskov & S. Molin, (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Applied and environmental microbiology* **64**: 2240-2246.

Andrews, S., (2014) FastQC: a quality control tool for high throughput sequence data. Version 0.11. 2. *Babraham Institute, Cambridge, UK* http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight & J.T. Eppig, (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**: 25-29.

Badouraly, R., M.C. Prevost, J.M. Ghigo & C. Beloin, (2010) Escherichia coli K-12 possesses multiple cryptic but functional chaperone–usher fimbriae with distinct surface specificities. *Environ. Microbiol.* **12**: 1957-1977.

Bagg, A. & J. Neilands, (1987) Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. *Biochemistry* **26**: 5471-5477.

Bak, G., J. Lee, S. Suk, D. Kim, J.Y. Lee, K.-s. Kim, B.-S. Choi & Y. Lee, (2015) Identification of novel sRNAs involved in biofilm formation, motility, and fimbriae formation in Escherichia coli. *Scientific reports* **5**.

Barrick, J.E., N. Sudarsan, Z. Weinberg, W.L. Ruzzo & R.R. Breaker, (2005) 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. *Rna* **11**: 774-784.

Barrios, A.F.G., R. Zuo, D. Ren & T.K. Wood, (2006) Hha, YbaJ, and OmpA regulate *Escherichia coli* K12 biofilm formation and conjugation plasmids abolish motility. *Biotechnology and Bioengineering* **93**: 188-200.

Beloin, C., A. Roux & J.-M. Ghigo, (2008) *Escherichia coli* biofilms. In: Bacterial Biofilms. Springer, pp. 249-289.

Bioinformatics, U.G., (2016) Frequently Asked Questions: Data File Formats. In., pp.

Blount, Z.D., (2015) The unexhausted potential of *E. coli*. *Elife* **4**: e05826.

Bolger, A.M., M. Lohse & B. Usadel, (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.

Bolstad, B.M., R.A. Irizarry, M. Åstrand & T.P. Speed, (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**: 185-193.

Boutte, C.C. & S. Crosson, (2013) Bacterial lifestyle shapes stringent response activation. *Trends in microbiology* **21**: 174-180.

Braun, V., (2003) Iron uptake by *Escherichia coli*. *Frontiers in Bioscience* **8**: s1409-s1421.

Bukau, B., (1993) Regulation of the *Escherichia coli* heat-shock response. *Molecular microbiology* **9**: 671-680.

Bullard, J.H., E. Purdom, K.D. Hansen & S. Dudoit, (2010) Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC bioinformatics* **11**: 94.

Bumgarner, R., (2013) Overview of DNA microarrays: types, applications, and their future. *Current protocols in molecular biology*: 22.21. 21-22.21. 11.

Cashel, M., (1996) The stringent response. *Escherichia coli and Salmonella typhimulium: cellular and molecular biology* **2**: 1458-1496.

Cavanagh, A.T., A.D. Klocko, X. Liu & K.M. Wassarman, (2008) Promoter specificity for 6S RNA regulation of transcription is determined by core promoter sequences and competition for region 4.2 of sigma70. *Molecular Microbiology* **67**: 1242-1256.

Chen, L., F. Sun, X. Yang, Y. Jin, M. Shi, L. Wang, Y. Shi, C. Zhan & Q. Wang, (2017) Correlation between RNA-Seq and microarrays results using TCGA data. *Gene* **628**: 200-204.

Cheung, K.J., V. Badarinarayana, D.W. Selinger, D. Janse & G.M. Church, (2003) A microarray-based antibiotic screen identifies a regulatory role for supercoiling in the osmotic stress response of *Escherichia coli*. *Genome research* **13**: 206-215.

Cho, B.-K., E.M. Knight & B.Ø. Palsson, (2006) Transcriptional regulation of the fad regulon genes of *Escherichia coli* by ArcA. *Microbiology* **152**: 2207-2219.

Christensen-Dalsgaard, M., M.G. Jorgensen & K. Gerdes, (2010a) Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Molecular Microbiology* **75**: 333-348.

Christensen-Dalsgaard, M., M.G. Jorgensen & K. Gerdes, (2010b) Three new RelE-homologous mRNA interferases of Escherichia coli differentially induced by environmental stresses. *Mol. Microbiol.* **75**: 333-348.

Conesa, A., P. Madrigal, S. Tarazona, D. Gomez-Cabrero, A. Cervera, A. McPherson, M.W. Szczesniak, D.J. Gaffney, L.L. Elo, X. Zhang & A. Mortazavi, (2016) A survey of best practices for RNA-seq data analysis. *Genome Biol* **17**: 13.

Costanzo, A. & S.E. Ades, (2006) Growth phase-dependent regulation of the extracytoplasmic stress factor, sigmaE, by guanosine 3',5'-bispyrophosphate (ppGpp). *Journal of Bacteriology* **188**: 4627-4634.

Cotter, P.A. & S. Stibitz, (2007) c-di-GMP-mediated regulation of virulence and biofilm formation. *Curr. Opin. Microbiol.* **10**: 17-23.

Croucher, N.J. & N.R. Thomson, (2010) Studying bacterial transcriptomes using RNA-seq. *Current opinion in microbiology* **13**: 619-624.

Dillies, M., A. Rau, J. Aubert, C. Hennequet-Antier, M. Jeanmougin, N. Servant, C. Keime, G. Marot, D. Castel & J. Estelle, (2012) On behalf of The French StatOmique Consortium: A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Briefings in Bioinformatics* **10**.

DiRusso, C.C. & T. Nyström, (1998) The fats of *Escherichia col*i during infancy and old age: regulation by global regulators, alarmones and lipid intermediates. *Molecular Microbiology* **27**: 1-8.

Domka, J., J. Lee & T.K. Wood, (2006) YliH (BssR) and YceP (BssS) regulate *Escherichia coli* K-12 biofilm formation by influencing cell signaling. *Applied and Environmental Microbiology* **72**: 2449-2459.

Dong, T. & H.E. Schellhorn, (2009a) Control of RpoS in global gene expression of *Escherichia coli* in minimal media. *Molecular Genetic and Genomics* **281**: 19-33.

Dong, T. & H.E. Schellhorn, (2009b) Global effect of RpoS on gene expression in pathogenic *Escherichia coli* O157:H7 strain EDL933. *BMC Genomics* **10**: 349.

Donlan, R.M. & J.W. Costerton, (2002) Biofilms: survival mechanisms of clinically relevant microorganisms. *Clinical microbiology reviews* **15**: 167-193.

Dorel, C., P. Lejeune & A. Rodrigue, (2006) The Cpx system of *Escherichia coli*, a strategic signaling pathway for confronting adverse conditions and for settling biofilm communities? *Research in microbiology* **157**: 306-314.

Dötsch, A., D. Eckweiler, M. Schniederjans, A. Zimmermann, V. Jensen, M. Scharfe, R. Geffers & S. Häussler, (2012) The Pseudomonas aeruginosa transcriptome in planktonic cultures and static biofilms using RNA sequencing. *PloS one* **7**: e31092.

Escolar, L., J. Pérez-Martín & V. De Lorenzo, (1999) Opening the iron box: transcriptional metalloregulation by the Fur protein. *Journal of bacteriology* **181**: 6223-6229.

Farewell, A., A.A. Diez, C.C. DiRusso & T. Nyström, (1996) Role of the *Escherichia coli* FadR regulator in stasis survival and growth phase-dependent expression of the *uspA, fad*, and *fab* genes. *Journal of bacteriology* **178**: 6443-6450.

Feng, Y. & J.E. Cronan, (2009) A new member of the *Escherichia coli fad* regulon: transcriptional regulation of *fadM* (*ybaW*). *Journal of Bacteriology* **191**: 6320-6328.

FLIN, K.P., (1987) The long-term survival of *Escherichia coli* in river water *Journal of Applied Bacteriology* **63**: 261-270.

Franchini, A.G. & T. Egli, (2006) Global gene expression in *Escherichia coli* K-12 during short-term and long-term adaptation to glucose-limited continuous culture conditions. *Microbiology* **152**: 2111-2127.

Franchini, A.G., J. Ihssen & T. Egli, (2015) Effect of global regulators RpoS and cyclic-AMP/CRP on the catabolome and transcriptome of *Escherichia coli* K12 during carbon- and energy-limited growth. *PLoS One* **10**: e0133793.

Giel, J.L., D. Rodionov, M. Liu, F.R. Blattner & P.J. Kiley, (2006) IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O2-regulated genes in *Escherichia coli*. *Molecular microbiology* **60**: 1058-1075.

Gonzalez Barrios, A.F., R. Zuo, Y. Hashimoto, L. Yang, W.E. Bentley & T.K. Wood, (2006) Autoinducer 2 controls biofilm formation in *Escherichia coli* through a novel motility quorum-sensing regulator (MqsR, B3022). *J Bacteriol* **188**: 305-316.

Goodall, E.C., A. Robinson, I.G. Johnston, S. Jabbari, K.A. Turner, A.F. Cunningham, P.A. Lund, J.A. Cole & I.R. Henderson, (2018) The essential genome of *Escherichia coli* K-12. *mBio* **9**: e02096-02017.

Graves, P.R. & T.A. Haystead, (2002) Molecular biologist's guide to proteomics. *Microbiology and Molecular Biology Reviews* **66**: 39-63.

Griffith, M., J.R. Walker, N.C. Spies, B.J. Ainscough & O.L. Griffith, (2015) Informatics for RNA sequencing: a web resource for analysis on the cloud. *PLoS computational biology* **11**: e1004393.

Han, Y., S. Gao, K. Muegge, W. Zhang & B. Zhou, (2015) Advanced applications of RNA sequencing and challenges. *Bioinformatics and biology insights* **9**: 29.

Hansen, A.M., Y. Qiu, N. Yeh, F.R. Blattner, T. Durfee & D.J. Jin, (2005) SspA is required for acid resistance in stationary phase by downregulation of H-NS in *Escherichia coli*. *Molecular microbiology* **56**: 719-734.

Harrison, J.J., W.D. Wade, S. Akierman, C. Vacchi-Suzzi, C.A. Stremick, R.J. Turner & H. Ceri, (2009) The chromosomal toxin gene yafQ is a determinant of multidrug tolerance for Escherichia coli growing in a biofilm. *Antimicrobial agents and chemotherapy* **53**: 2253-2258.

Heather, J.M. & B. Chain, (2016) The sequence of sequencers: The history of sequencing DNA. *Genomics* **107**: 1-8.

Henderson, I.R. & P. Owen, (1999) The Major Phase-Variable Outer Membrane Protein ofEscherichia coli Structurally Resembles the Immunoglobulin A1 Protease Class of Exported Protein and Is Regulated by a Novel Mechanism Involving Dam and OxyR. *J. Bacteriol.* **181**: 2132-2141.

Hengge-Aronis, R., W. Klein, R. Lange, M. Rimmele & W. Boos, (1991) Trehalose synthesis genes are controlled by the putative sigma factor encoded by rpoS and are involved in stationary-phase thermotolerance in *Escherichia coli*. *Journal of Bacteriology* **173**: 7918-7924.

Hengge-Aronis, R.L.a.R., (1991a) Growth phase-regulated expression of *bolA* and morphology of stationary-phase *Escherichia coli* cells are controlled by the novel sigma factor RpoS. *Journal of Bacteriology*: 4474-4481.

Hengge-Aronis, R.L.a.R., (1991b) Identification of a central regulator of stationary-phase gene expression in *Escherichia* coii. *Motecuiar Microbiology* **5**: 49-59.

Hofmann, N., R. Wurm & R. Wagner, (2011) The *E. coli* anti-sigma factor Rsd: studies on the specificity and regulation of its expression. *PLoS One* **6**: e19235.

Hsiao, A. & J. Zhu, (2009) Genetic tools to study gene expression during bacterial pathogen infection. *Advances in applied microbiology* **67**: 297-314.

Hu, M. & K. Polyak, (2006) Serial analysis of gene expression. *Nature protocols* **1**: 1743.

Hu, Y., B.W. Kwan, D.O. Osbourne, M.J. Benedik & T.K. Wood, (2015) Toxin YafQ increases persister cell formation by reducing indole signalling. *Environmental microbiology* **17**: 1275-1285.

Huai-Shu XU, N.R., F. L. Singleton, R. W. Attwell, D. J. Grimes, & a.R.R. ColwelP, (1982) Survival and Viability of Nonculturable *Escherichia coli* and Vibrio cholerae in the Estuarine and Marine Environment *Microb. Ecol* **8**: 313-323.

Hubbell, E., W.-M. Liu & R. Mei, (2002) Robust estimators for expression analysis. *Bioinformatics* **18**: 1585-1592.

Inouye, K.y.a.M., (1997) Growth-Phase-Dependent Expression of cspD, Encoding a Member of the CspA Family in *Escherichia coli*. *JOURNAL OF BACTERIOLOGY*: 5126-5130.

Irizarry, R.A., B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf & T.P. Speed, (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**: 249-264.

Ishihama, A., (2000) Functional modulation of *Escherichia coli* RNA polymerase. *Annual Reviews in Microbiology* **54**: 499-518.

Ishihama, M.J.a.A., (1998) A stationary phase protein in *Escherichia coli* with binding activity to the major subunit of RNA polymerase. *PNAS* **95**: 4953–4958.

Itoh, Y., J.D. Rice, C. Goller, A. Pannuri, J. Taylor, J. Meisner, T.J. Beveridge, J.F. Preston & T. Romeo, (2008) Roles of pgaABCD genes in synthesis, modification, and export of the Escherichia coli biofilm adhesin poly-β-1, 6-N-acetyl-D-glucosamine. *J. Bacteriol.* **190**: 3670-3680.

Jaluria, P., K. Konstantopoulos, M. Betenbaugh & J. Shiloach, (2007) A perspective on microarrays: current applications, pitfalls, and potential uses. *Microbial cell factories* **6**: 4.

Jayaraman, R., (2008) Bacterial persistence: some new insights into an old phenomenon. *Journal of biosciences* **33**: 795-805.

Jonas, K., A.N. Edwards, R. Simm, T. Romeo, U. Römling & Ö. Melefors, (2008) The RNA binding protein CsrA controls cyclic di-GMP metabolism by directly regulating the expression of GGDEF proteins. *Mol. Microbiol.* **70**: 236-257.

Kabir, M.S., D. Yamashita, S. Koyama, T. Oshima, K. Kurokawa, M. Maeda, R. Tsunedomi, M. Murata, C. Wada & H. Mori, (2005) Cell lysis directed by σE in early stationary phase and effect of induction of the rpoE gene on global gene expression in *Escherichia coli*. *Microbiology* **151**: 2721-2735.

Kannan, G., J.C. Wilks, D.M. Fitzgerald, B.D. Jones, S.S. BonDurant & J.L. Slonczewski, (2008) Rapid acid treatment of *Escherichia coli* : transcriptomic response and recovery. *BMC Microbiology* **8**: 37.

Karp, P.D., D. Weaver, S. Paley, C. Fulcher, A. Kubo, A. Kothari, M. Krummenacker, P. Subhraveti, D. Weerasinghe, S. Gama-Castro, A.M. Huerta, L. Muniz-Rascado, C. Bonavides-Martinez, V. Weiss, M. Peralta-Gil, A. Santos-Zavaleta, I. Schroder, A. Mackie, R. Gunsalus, J. Collado-Vides, I.M. Keseler & I. Paulsen, (2014) The EcoCyc Database. *EcoSal Plus* **6**.

Karpinets, T.V., D.J. Greenwood, C.E. Sams & J.T. Ammons, (2006) RNA: protein ratio of the unicellular organism as a characteristic of phosphorous and nitrogen stoichiometry and of the cellular requirement of ribosomes for protein synthesis. *BMC biology* **4**: 30.

Kasari, V., K. Kurg, T. Margus, T. Tenson & N. Kaldalu, (2010) The *Escherichia coli* mqsR and ygiT genes encode a new toxin-antitoxin pair. *Journal of bacteriology* **192**: 2908-2919.

Keren, I., D. Shah, A. Spoering, N. Kaldalu & K. Lewis, (2004) Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *Journal of bacteriology* **186**: 8172-8180.

Kim, Y., X. Wang, X.S. Zhang, S. Grigoriu, R. Page, W. Peti & T.K. Wood, (2010) *Escherichia coli* toxin/antitoxin pair MqsR/MqsA regulate toxin CspD. *Environmental microbiology* **12**: 1105-1121.

Kim, Y. & T.K. Wood, (2010) Toxins Hha and CspD and small RNA regulator Hfq are involved in persister cell formation through MqsR in *Escherichia coli*. *Biochemical and biophysical research communications* **391**: 209-213.

Kirawada, Y., Nobuyukifujitat,and Akiraishihama, (1990) Structure and probable genetic location of a "ribosome modulation factor" associated with 100S ribosomes in stationary-phase Escherichiacoli cells. *PNAS* **87**: 2657-2661.

Kiupakis, A.K. & L. Reitzer, (2002) ArgR-independent induction and ArgR-dependent superinduction of the astCADBE operon in *Escherichia coli*. *Journal of bacteriology* **184**: 2940-2950.

Kjaergaard K, S.M., Hasman H, Klemm P., (2000) Antigen 43 from *Escherichia coli* induces inter- and intraspecies cell aggregation and changes in colony morphology of *Pseudomonas fluorescens*. *Journal of Applied Bacteriology* **182**: 4789-4796.

Kolodkin-Gal, I., R. Verdiger, A. Shlosberg-Fedida & H. Engelberg-Kulka, (2009) A differential effect of *E. coli* toxin-antitoxin systems on cell death in liquid media and biofilm formation. *PLoS One* **4**: e6785.

Kostić, T., A. Weilharter, S. Rubino, G. Delogu, S. Uzzau, K. Rudi, A. Sessitsch & L. Bodrossy, (2007) A microbial diagnostic microarray technique for the sensitive detection and identification of pathogenic bacteria in a background of nonpathogens. *Analytical biochemistry* **360**: 244-254.

Kumari, S., C.M. Beatty, D.F. Browning, S.J. Busby, E.J. Simel, G. Hovel-Miner & A.J. Wolfe, (2000) Regulation of acetyl coenzyme A synthetase inEscherichia coli. *Journal of bacteriology* **182**: 4173-4179.

Lacour, S. & P. Landini, (2004) SigmaS-dependent gene expression at the onset of stationary phase in *Escherichia coli*: function of sigmaS-dependent genes and identification of their promoter sequences. *J Bacteriol* **186**: 7186-7195.

Lal, A., S. Krishna & A.S.N. Seshasayee, (2016) Regulation of global transcription in *E. coli* by Rsd and 6S RNA. *bioRxiv*: 058339.

Landini, P., (2009) Cross-talk mechanisms in biofilm formation and responses to environmental and physiological stress in *Escherichia coli*. *Research in microbiology* **160**: 259-266.

Langmead, B. & S.L. Salzberg, (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**: 357-359.

Larsonneur, F., F.A. Martin, A. Mallet, M. Martinez-Gil, V. Semetey, J.M. Ghigo & C. Beloin, (2016) Functional analysis of *Escherichia coli* Yad fimbriae reveals their potential role in environmental persistence. *Environ Microbiol* **18**: 5228-5248.

Lee, J., S. Hiibel, K. Reardon & T. Wood, (2010) Identification of stress-related proteins in *Escherichia coli* using the pollutant cis-dichloroethylene. *Journal of applied microbiology* **108**: 2088-2102.

Lee, J., R. Page, R. García-Contreras, J.-M. Palermino, X.-S. Zhang, O. Doshi, T.K. Wood & W. Peti, (2007) Structure and function of the *Escherichia coli* protein YmgB: a protein critical for biofilm formation and acid-resistance. *Journal of molecular biology* **373**: 11-26.

Lee, S.K., J.D. Newman & J.D. Keasling, (2005) Catabolite repression of the propionate catabolic genes in *Escherichia coli* and Salmonella enterica: evidence for involvement of the cyclic AMP receptor protein. *Journal of Bacteriology* **187**: 2793-2800.

Lee, S.Y., (1996) High cell-density culture of *Escherichia coli*. *Trends Biotechnology* **4**: 98-105.

Li, C. & W.H. Wong, (2001) Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. *Genome biology* **2**: research0032. 0031.

Liao, Y., G.K. Smyth & W. Shi, (2013) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923-930.

Lim, W.K., K. Wang, C. Lefebvre & A. Califano, (2007) Comparative analysis of microarray normalization procedures: effects on reverse engineering gene networks. *Bioinformatics* **23**: i282-i288.

Lin, T., E.B. Troy, L.T. Hu, L. Gao & S.J. Norris, (2014) Transposon mutagenesis as an approach to improved understanding of *Borrelia* pathogenesis and biology. *Frontiers in Cellular and Infection Microbiology* **4**: 63.

Loots, G.G., (2008) Genomic identification of regulatory elements by evolutionary sequence comparison and functional analysis. *Advances in genetics* **61**: 269-293.

Love, M., S. Anders & W. Huber, (2014) Differential analysis of count data–the DESeq2 package. *Genome Biol* **15**: 550.

Makinoshima, H., S.I. Aizawa, H. Hayashi, T. Miki, A. Nishimura & A. Ishihama, (2003) Growth Phase-Coupled Alterations in Cell Structure and Function of *Escherichia coli*. *Journal of Bacteriology* **185**: 1338-1345.

Martin, J.E., L.S. Waters, G. Storz & J.A. Imlay, (2015) The *Escherichia coli* small protein MntS and exporter MntP optimize the intracellular concentration of manganese. *PLoS genetics* **11**: e1004977.

Masahiro Yamagishi, H.M., Akira Wada1, Masayuki Sakagami, Nobuyuki Fujita and Akira Ishihama, (1993) Regulation of the *Escherichia coli* rmf gene encoding the ribosome modulation factor: growth phase- and growth rate-dependent control. *EMBO J* **12**: 625-630.

McGrath, K.C., S.R. Thomas-Hall, C.T. Cheng, L. Leo, A. Alexa, S. Schmidt & P.M. Schenk, (2008) Isolation and analysis of mRNA from environmental microbial communities. *Journal of Microbiological Methods* **75**: 172-176.

McHugh, J.P., F. Rodríguez-Quiñones, H. Abdul-Tehrani, D.A. Svistunenko, R.K. Poole, C.E. Cooper & S.C. Andrews, (2003) Global iron-dependent gene regulation in *Escherichia coli* : a new mechanism for iron homeostasis. *Journal of Biological Chemistry*.

Metzner, M., J. Germer & R. Hengge, (2004) Multiple stress signal integration in the regulation of the complex σS-dependent csiD-ygaF-gabDTP operon in *Escherichia coli*. *Molecular microbiology* **51**: 799-811.

Michal Aviv, H.G., Amos B. Oppenheim, Gad Glaser (1996) Analysis of the shut-off of ribosomal RNA promoters in *Escherichia coli* upon entering the stationary phase of growth. *FEMS Microbiol Lett* **140**: 71-76.

Miller, M.B. & Y.-W. Tang, (2009) Basic concepts of microarrays and potential applications in clinical microbiology. *Clinical microbiology reviews* **22**: 611-633.

Mitchell, J.E., T. Oshima, S.E. Piper, C.L. Webster, L.F. Westblade, G. Karimova, D. Ladant, A. Kolb, J.L. Hobman & S.J. Busby, (2007) The *Escherichia coli* regulator of sigma 70 protein, Rsd, can up-regulate some stress-dependent promoters by sequestering sigma 70. *Journal of Bacteriology* **189**: 3489-3495.

Monje-Casas, F., J. Jurado, M.a.-J. Prieto-Álamo, A. Holmgren & C. Pueyo, (2001) Expression Analysis of the nrdHIEF Operon fromEscherichia coli CONDITIONS THAT TRIGGER THE TRANSCRIPT LEVEL IN VIVO. *Journal of Biological Chemistry* **276**: 18031-18037.

Mortazavi, A., B.A. Williams, K. McCue, L. Schaeffer & B. Wold, (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* **5**: 621.

N. Gustavsson, A.A.D.a.T.N.m., (2002) The universal stress protein paralogues of *Escherichia coli* are co-ordinately regulated and co-operate in the defence against DNA damage. *Molecular Microbiology* **42**: 107-117.

Naef, F. & M.O. Magnasco, (2003) Solving the riddle of the bright mismatches: labeling and effective binding in oligonucleotide arrays. *Physical Review E* **68**: 011906.

Nagalakshmi, U., K. Waern & M. Snyder, (2010) RNA-Seq: a method for comprehensive transcriptome analysis. *Current Protocols in Molecular Biology*: 4.11. 11-14.11. 13.

Neusser, T., T. Polen, R. Geissen & R. Wagner, (2010) Depletion of the non-coding regulatory 6S RNA in *E. coli* causes a surprising reduction in the expression of the translation machinery. *BMC Genomics* **11**: 165.

Niven, G.W. & W.M. El-Sharoud, (2008) Ribosome modulation factor. In: Bacterial Physiology. Springer, pp. 293-311.

Nonaka, G., M. Blankschien, C. Herman, C.A. Gross & V.A. Rhodius, (2006) Regulon and promoter analysis of the *E. coli* heat-shock factor, σ32, reveals a multifaceted cellular response to heat stress. *Genes & Development* **20**: 1776-1789.

Nookaew, I., M. Papini, N. Pornputtapong, G. Scalcinati, L. Fagerberg, M. Uhlén & J. Nielsen, (2012) A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in Saccharomyces cerevisiae. *Nucleic Acids Research* **40**: 10084-10097.

Nuccio, S.-P. & A.J. Bäumler, (2007) Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol. Mol. Biol. Rev.* **71**: 551-575.

Pachter, L., (2011) Models for transcript quantification from RNA-Seq. *arXiv preprint arXiv:1104.3889*.

Patten, C.L., M.G. Kirchhof, M.R. Schertzberg, R.A. Morton & H.E. Schellhorn, (2004) Microarray analysis of RpoS-mediated gene expression in *Escherichia coli* K-12. *Mol Genet Genomics* **272**: 580-591.

Penders, J., C. Thijs, C. Vink, F.F. Stelma, B. Snijders, I. Kummeling, P.A. van den Brandt & E.E. Stobberingh, (2006) Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* **118**: 511-521.

Pepke, S., B. Wold & A. Mortazavi, (2009) Computation for ChIP-seq and RNA-seq studies. *Nature methods* **6**: S22-S32.

Pesavento, C., G. Becker, N. Sommerfeldt, A. Possling, N. Tschowri, A. Mehlis & R. Hengge, (2008) Inverse regulatory coordination of motility and curli-mediated adhesion in *Escherichia coli*. *Genes & development* **22**: 2434-2446.

Petrova, O.E., F. Garcia-Alcalde, C. Zampaloni & K. Sauer, (2017) Comparative evaluation of rRNA depletion procedures for the improved analysis of bacterial biofilm and mixed pathogen culture transcriptomes. *Scientific reports* **7**: 41114.

Piper, S.E., J.E. Mitchell, D.J. Lee & S.J. Busby, (2009) A global view of *Escherichia coli* Rsd protein and its interactions. *Molecular BioSystems* **5**: 1943-1947.

Povolotsky, T.L. & R. Hengge, (2012) 'Life-style'control networks in *Escherichia coli*: signaling by the second messenger c-di-GMP. *Journal of Biotechnology* **160**: 10-16.

Prysak, M.H., C.J. Mozdzierz, A.M. Cook, L. Zhu, Y. Zhang, M. Inouye & N.A. Woychik, (2009) Bacterial toxin YafQ is an endoribonuclease that associates with the ribosome and blocks translation elongation through sequence-specific and frame-dependent mRNA cleavage. *Molecular microbiology* **71**: 1071-1087.

Reimers, M. & V.J. Carey, (2006) [8] Bioconductor: an open source framework for bioinformatics and computational biology. *Methods in enzymology* **411**: 119-134.

Ren, D., L.A. Bedzyk, S.M. Thomas, R.W. Ye & T.K. Wood, (2004) Gene expression in *Escherichia coli* biofilms. *Appl Microbiol Biotechnol* **64**: 515-524.

Renilla, S., V. Bernal, T. Fuhrer, S. Castaño-Cerezo, J.M. Pastor, J.L. Iborra, U. Sauer & M. Cánovas, (2012) Acetate scavenging activity in *Escherichia coli*: interplay of acetyl–

CoA synthetase and the PEP–glyoxylate cycle in chemostat cultures. *Applied Microbiology and Biotechnology* **93**: 2109-2124.

Robinson, M.D. & A. Oshlack, (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology* **11**: R25.

Roux, A., C. Beloin & J.-M. Ghigo, (2005) Combined inactivation and expression strategy to study gene function under physiological conditions: application to identification of new Escherichia coli adhesins. *J. Bacteriol.* **187**: 1001-1013.

Ryan, M.M., S.J. Huffaker, M.J. Webster, M. Wayland, T. Freeman & S. Bahn, (2004) Application and optimization of microarray technologies for human postmortem brain studies. *Biological psychiatry* **55**: 329-336.

Schellhorn, H., (1995) Regulation of hydroperoxidase (catalase) expression in *Escherichia coli*. *FEMS microbiology letters* **131**: 113-119.

Schellhorn HE, A.J., Wei LI, Chang L., (1998) Identification of conserved, RpoS-dependent stationary-phase genes of Escherichia coli. *Journal of bacteriology* **180**: 6283-6291.

Schellhorn, H.E., J.P. Audia, L.I. Wei & L. Chang, (1998) Identification of conserved, RpoS-dependent stationary-phase genes of *Escherichia coli*. *Journal of Bacteriology* **180**: 6283-6291.

Schembri, M.A., K. Kjærgaard & P. Klemm, (2003) Global gene expression in *Escherichia coli* biofilms. *Molecular Microbiology* **48**: 253–267.

Schneider, B.L., A.K. Kiupakis & L.J. Reitzer, (1998) Arginine catabolism and the arginine succinyltransferase pathway in *Escherichia coli*. *Journal of bacteriology* **180**: 4278-4286.

Schulze, A. & J. Downward, (2001) Navigating gene expression using microarrays—a technology review. *Nature cell biology* **3**: E190-E195.

Schuster, E.F., E. Blanc, L. Partridge & J.M. Thornton, (2007) Estimation and correction of non-specific binding in a large-scale spike-in experiment. *Genome Biology* **8**: R126.

Seo, S.W., D. Kim, H. Latif, E.J. O'Brien, R. Szubin & B.O. Palsson, (2014) Deciphering Fur transcriptional regulatory network highlights its complex role beyond iron metabolism in *Escherichia coli*. *Nature communications* **5**: 4910.

Sezonov, G., D. Joseleau-Petit & R. D'Ari, (2007) *Escherichia coli* physiology in Luria-Bertani broth. *J Bacteriol* **189**: 8746-8749.

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

Sheikh, A., Q. Luo, K. Roy, S. Shabaan, P. Kumar, F. Qadri & J.M. Fleckenstein, (2014) Contribution of the highly conserved EaeH surface protein to enterotoxigenic Escherichia coli pathogenesis. *Infect. Immun.* **82**: 3657-3666.

Shendure, J. & E.L. Aiden, (2012) The expanding scope of DNA sequencing. *Nature biotechnology* **30**: 1084.

Shi, Y., G.W. Tyson & E.F. DeLong, (2009) Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature* **459**: 266.

Smyth, G.K., (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology* **3**: 1-25.

Smyth, G.K. & T. Speed, (2003) Normalization of cDNA microarray data. *Methods* **31**: 265-273.

Sohanpal, B.K., S. El-Labany, M. Lahooti, J.A. Plumbridge & I.C. Blomfield, (2004) Integrated regulatory responses of fimB to N-acetylneuraminic (sialic) acid and GlcNAc in Escherichia coli K-12. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 16322-16327.

Sommerfeldt, N., A. Possling, G. Becker, C. Pesavento, N. Tschowri & R. Hengge, (2009) Gene expression patterns and differential input into curli fimbriae regulation of all GGDEF/EAL domain proteins in *Escherichia coli*. *Microbiology* **155**: 1318-1331.

Su, C. & L.M. Sordillo, (1998) A simple method to enrich mRNA from total prokaryotic RNA. *Molecular biotechnology* **10**: 83-85.

Takahashi, Y. & U. Tokumoto, (2002) A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids. *Journal of Biological Chemistry* **277**: 28380-28383.

Thomas Neusser, T.P., René Geissen, Rolf Wagner, (2010) Depletion of the non-coding regulatory 6S RNA in *E. coli* causes a surprising reduction in the expression of the translation machinery. *BMC Genomics* **11**.

Uliczka, F., F. Pisano, A. Kochut, W. Opitz, K. Herbst, T. Stolz & P. Dersch, (2011) Monitoring of gene expression in bacteria during infections using an adaptable set of bioluminescent, fluorescent and colorigenic fusion vectors. *PLoS One* **6**: e20425.

Valgepea, K., K. Adamberg, R. Nahku, P.-J. Lahtvee, L. Arike & R. Vilu, (2010) Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase. *BMC systems biology* **4**: 166.

Van Opijnen, T., K.L. Bodi & A. Camilli, (2009) Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature methods* **6**: 767.

Van Vliet, A.H., (2010) Next generation sequencing of microbial transcriptomes: challenges and opportunities. *FEMS microbiology letters* **302**: 1-7.

Vega, N.M., K.R. Allison, A.S. Khalil & J.J. Collins, (2012) Signaling-mediated bacterial persister formation. *Nature chemical biology* **8**: 431-433.

Verhaak, R.G., F.J. Staal, P.J. Valk, B. Lowenberg, M.J. Reinders & D. de Ridder, (2006) The effect of oligonucleotide microarray data pre-processing on the analysis of patient-cohort studies. *BMC bioinformatics* **7**: 105.

Vogeleer, P., Y.D. Tremblay, A.A. Mafu, M. Jacques & J. Harel, (2014) Life on the outside: role of biofilms in environmental persistence of Shiga-toxin producing *Escherichia coli*. *Front Microbiol* **5**: 317.

Wada, A., K. Igarashi, S. Yoshimura, S. Aimoto & A. Ishihama, (1995) Ribosome modulation factor: stationary growth phase-specific inhibitor of ribosome functions from *Escherichia coli*. *Biochemical and biophysical research communications* **214**: 410-417.

Wang, X., Y. Kim, S.H. Hong, Q. Ma, B.L. Brown, M. Pu, A.M. Tarone, M.J. Benedik, W. Peti & R. Page, (2011) Antitoxin MqsA helps mediate the bacterial general stress response. *Nature chemical biology* **7**: 359-366.

Wang, X. & T.K. Wood, (2011) Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. *Appl Environ Microbiol* **77**: 5577-5583.

Wang, Z., M. Gerstein & M. Snyder, (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews genetics* **10**: 57-63.

Wassarman, K.M. & G. Storz, (2000) 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* **101**: 613-623.

Weber, H., T. Polen, J. Heuveling, V.F. Wendisch & R. Hengge, (2005) Genome-wide analysis of the general stress response network in *Escherichia coli*: sigmaS-dependent genes, promoters, and sigma factor selectivity. *J Bacteriol* **187**: 1591-1603.

Wendisch, V.F., D.P. Zimmer, A. Khodursky, B. Peter, N. Cozzarelli & S. Kustu, (2001) Isolation of *Escherichia coli* mRNA and comparison of expression using mRNA and total RNA on DNA microarrays. *Analytical biochemistry* **290**: 205-213.

Williams, M.D., T.X. Ouyang & M.C. Flickinger, (1994) Starvation-induced expression of SspA and SspB: the effects of a null mutation in sspA on *Escherichia coli* protein synthesis and survival during growth and prolonged starvation. *Molecular microbiology* **11**: 1029-1043.

Winfield, M.D. & E.A. Groisman, (2003) Role of Nonhost Environments in the Lifestyles of Salmonella and *Escherichia coli*. *Applied and Environmental Microbiology* **69**: 3687-3694.

Winkelmann, G., (2002) Microbial siderophore-mediated transport. In.: Portland Press Limited, pp.

Wong, G.T., R.P. Bonocora, A.N. Schep, S.M. Beeler, A.J.L. Fong, L.M. Shull, L.E. Batachari, M. Dillon, C. Evans & C.J. Becker, (2017) Genome-wide transcriptional response to varying RpoS levels in Escherichia coli K-12. *Journal of bacteriology* **199**: e00755-00716.

Wood, T.K., S.J. Knabel & B.W. Kwan, (2013) Bacterial persister cell formation and dormancy. *Appl Environ Microbiol* **79**: 7116-7121.

Wu, Y. & F.W. Outten, (2009) IscR controls iron-dependent biofilm formation in *Escherichia coli* by regulating type I fimbria expression. *Journal of bacteriology* **191**: 1248-1257.

Xing, Y., K. Kapur & W.H. Wong, (2006) Probe selection and expression index computation of affymetrix exon arrays. *PloS one* **1**: e88.

Zhai, Y. & M.H. Saier, (2002) The β-barrel finder (BBF) program, allowing identification of outer membrane β-barrel proteins encoded within prokaryotic genomes. *Protein Sci.* **11**: 2196-2207.

Zhang, X.S., R. Garcia-Contreras & T.K. Wood, (2008) *Escherichia coli* transcription factor YncC (McbR) regulates colanic acid and biofilm formation by repressing expression of periplasmic protein YbiM (McbA). *ISME J* **2**: 615-631.

Zhao, S., Fung-Leung, Wai-Ping, A. Bittner, K. Ngo & X. Liu, (2014) Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PloS one* **9**: e78644.

Zheng, M., B. Doan, T.D. Schneider & G. Storz, (1999) OxyR and SoxRS regulation of *fur*. *Journal of bacteriology* **181**: 4639-4643.

Zhou, J., (2003) Microarrays for bacterial detection and microbial community analysis. *Current opinion in Microbiology* **6**: 288-294.