Transcriptional analysis of *Escherichia coli* O157:H7 during *in vivo* mimicking conditions

by

Michael D. Carruthers

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Microbiology

Program of Study Committee:
F. Chris Minion, Major Professor
Jeff Beetham
Nancy Cornick
Larry Halverson
Dan Nordman

Iowa State University
Ames, Iowa
2009

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abstract

Most *Escherichia coli* reside harmlessly in the animal intestine and share a mutualistic interaction with its human host. This relationship endures until the scales are shifted toward the pathogen such as a break in the intestinal mucosal barrier, host immunosuppression, the introduction of the species into a non-native environment such as the urinary tract, acquisition of a new phenotype, appearance of a strain with the ability to cause disease in a healthy host, etc. The focus of this dissertation is the interaction of the pathogen *E. coli* O157:H7 with the ruminant host. Four separate microarray studies were conducted to ascertain the transcriptome of *E. coli* O157:H7 under environmental conditions found within the ruminant host. These conditions were heat shock, growth under anaerobic conditions, exposure to rumen fluid, and interaction with a protozoan.

The response to elevated temperature (heat shock) is highly conserved. The transcriptome of *Escherichia coli* K-12 has been studied over a variety of conditions while such studies involving *E. coli* O157:H7 are just now being conducted. To better understand the impact of heat shock on *E. coli* O157:H7, global transcript levels of strain EDL933 cells shifted from 37°C to 50°C for 15 min were compared to cells left at 37°C by microarray. Using a mixed model analysis, 193 genes were found to be differentially transcribed at $p < 0.0042$ with a q-value of < 0.1. The 111 down-regulated genes include the curli pili associated genes *csgABCDEFG,* maltose transport associated proteins *malEFK,* and NADH dehydrogenase subunit encoding *nuoCEHIJN*. The 82 genes up-regulated include the heat shock induced genes *rpoH*, *dnaK*, *dnaJ*, *groEL*, *groES*, and *grpE* along with two LEE encoded genes; hypothetical gene Z5121 and *sepZ*. Twenty-three additional genes located in O-islands were found to be differentially expressed. qRT-PCR was performed to validate microarray results. Also, samples subjected to a 30˚C to 42˚C shift were examined by qRT-PCR to confirm differential transcription of selected genes. These results indicate that this pathogen may regulate its virulence factors in response to temperature changes.

Escherichia coli O157:H7 encounters two environmental conditions within the ruminant gastrointestinal tract, limiting oxygen and rumen fluid. Both of these conditions may have implications in the general physiology of *E. coli* and its survival during passage through the rumen. The impact of these environmental conditions on the transcriptome of *E. coli* O157:H7 is unclear. To study this further, the transcriptome of *E. coli* O157:H7 strain EDL933 grown under anaerobic conditions was compared to cells exposed to rumen fluid and to cells grown aerobically. Four hundred nineteen genes were differentially expressed during anaerobic growth at $p < 0.014$ with a false discovery rate < 10%. Among the 280 up-regulated genes were the curli pili associated genes *csgBA and csgDEFG*, adherence associated genes *ompA, tdcA,* and *cadA,* heat shock genes *dnaK*, *dnaJ*, *groEL*, and *groES* and the σ^{38} encoding gene *rpoS*. Ten genes, 8 up- and 2 down-regulated, located in *E. coli* O157:H7 virulence-related O-islands were differentially transcribed with a fold change >1.3 suggesting that these genes may play a role in host adaptation. Translational β-galactosidase fusions with *csgBA* confirmed its regulation during anaerobic conditions. As a result of rumen fluid exposure, 83 genes were differentially transcribed at $p < 0.002$ with a false discovery rate < 10%, 38 up- and 45 down-regulated. There were 11 up-regulated, 11 down-regulated, and 3 opposite-regulated genes common to both conditions. Our analysis indicated that protein synthesis was retarded during anaerobiasis and may be further slowed by exposure to rumen fluid.

Free-living protozoa, such as *Acanthamoeba castellanii*, are environmental hosts for several bacterial species pathogenic to humans and animals. Intracellular residence within protozoa has been implicated in enhancing virulence by selecting for virulence factors and antibiotic resistance outside of the animal host. To better understand this relationship with *E. coli* O157:H7, we characterized *E. coli* O157:H7 transcriptomes from extracellular bacteria in grown in broth to *E. coli* O157:H7 residing within *A. castellanii* using two-color microarrays. Statistical analysis indicated that 969 genes were differentially expressed at a p value < 0.018, with a false discovery rate of 1.9% and a fold change cutoff of 1.3. There were 655 up-regulated transcripts that include 40 genes involved in virulence of which 32 are encoded on O-islands. These up-regulated

virulence-associated genes included shiga toxin genes (*stx*1A, *stx*1B *stx*2A), 14 genes involved in LEE and non-LEE encoded type three secretion system components and secreted factors. Additional induced genes include SOS response genes such as *lexA* and *recA*, genes involved in or predicted to be involved in antibiotic resistance (*rarD, macAB, marABR, mdtK, yojI, yhgN*), the quorum sensing operon *lsrACDB*, the *efe* and *feo* iron acquisition systems There were 314 down-regulated transcripts that include 19 transcripts associated with virulence, 7 of which are encoded on O-islands. Our results demonstrate a significant portion of the *E. coli* O157:H7 genome, including many virulence related genes, was differentially expressed as a result of the protozoan intracellular environment.

In summary, these studies attempt to shed light on the biology of *E. coli* O157:H7 within the ruminant host. These results show that *E. coli* O157:H7 demonstrates gene regulation similar to that of non-pathogenic *E. coli* and other bacteria in response to the aforementioned environments. In addition, *E. coli* O157:H7 demonstrated differential expression of virulence factors and genes found within other pathogenic *E. coli* along with genes that have not been characterized as being differentially transcribed or associated with the previously described environments. The characterization of *E. coli* O157:H7 within these environments has provided insights into unanswered questions and on unknown mechanisms of survival and persistence of this important pathogen.

CHAPTER 1. GENERAL INTRODUCTION

Introduction

Escherichia coli is a small constituent of the global micro-biosphere being mainly relegated to the gastrointestinal tracts of warm-blooded mammals and excretions thereof. Even though it is a minority in terms of spatial distribution and shear numbers, a disproportionate amount of research focuses on or involves *E. coli* in some way. There are many reasons for this. Simple growth requirements coupled with a short generation time allows for a large amount of biomass to be generated cheaply in a short period of time. Genetic manipulation of this species has proven to be relatively easy; it can be transformed via artificial competence, transduction, or conjugation. *Escherichia coli* is also capable of harboring a variety of plasmids and transposons allowing for more opportunities for genetic analyses. As knowledge grew, *E. coli* became an increasingly attractive experimental system. In addition, the species has numerous strains of medical importance, which has driven much of the current research.

Most *E. coli* reside harmlessly in the animal intestine and share a mutualistic interaction with its human host. This relationship endures until the scales are shifted toward the pathogen such as a break in the intestinal mucosal barrier, host immunosuppression, the introduction of the species into a non-native environment such as the urinary tract, acquisition of a new phenotype, appearance of a strain with the ability to cause disease in a healthy host, etc. The last relationship described above is the focus of this dissertation, specifically the interaction of *E. coli* O157:H7 with the ruminant host.

The work outlined in the following studies will aid in an increase of knowledge regarding *E. coli* O157:H7 and the extent to which it regulates transcription in response to environmental conditions found in the ruminant gastrointestinal tract as it proceeds from the mouth through the anus.

Dissertation Organization

This dissertation is organized into five chapters, including a General Introduction, three chapters to be submitted for publication, and a final Summary and Discussion with suggested directions for future studies. Chapter 1 is an introduction containing a literature review of enterohemorrhagic *E. coli*, specifically *E. coli* O157:H7, and microarrays. Chapter 2 describes a study comparing transcriptional changes in *E. coli* O157:H7 during heat shock that has been published in FEMS Micobiol. Lett. (2009, 295, 96-102). Chapter 3 examines transcriptional differences between *E. coli* O157:H7 cells grown under anaerobic conditions compared to aerobic conditions and exposure to rumen fluid anaerobically. Transcriptional differences between *E. coli* O157:H7 within protozoa compared to planktonic cells are discussed in Chapter 4. Conclusions and ideas for future research are discussed in Chapter 5. This dissertation uses the reference format for Infection and Immunity throughout except for Chapter 2, which has been accepted for publication in FEMS Microbiology Letters. Chapter 2 uses that journal's format for its references. There are four appendices; Appendix A lists the differentially expressed genes of *E. coli* O157:H7 as a result of heat shock; Appendix B lists the differentially expressed genes of *E. coli* O157:H7 as a result of anaerobic growth; Appendix C lists the differentially expressed genes of *E. coli* O157:H7 as a result of rumen fluid treatment; and Appendix D lists the differentially expressed genes of *E. coli* O157:H7 within *Acanthamoeba castellanii*.

Literature Review

Escherichia coli

In 1885 Theodore Escherich described a pure bacterial culture of organisms with slim, short rods that was later named *Escherichia coli.* This microorganism colonizes the intestinal tract of warm-blooded animals within a few hours after birth and has been shown to play an important role in maintaining intestinal physiology (10). *Escherichia coli* are thought to play at least two beneficial roles within the human intestine such as stimulating immune function and aiding in competitive exclusion of harmful bacteria. These bacteria also have harmful roles such as the existence of pathogenic strains, production of potential carcinogens and intestinal putrefaction (35, 49).

Pathogenic *Escherichia coli*

Pathogenic *E. coli* have evolved from commensals through the acquisition of virulence factors by multiple means of genetic exchange (bacteriophage, plasmids, and transposons). Diarrheagenic *E. coli* can be categorized into viro/pathotypes based on virulence attributes, mechanisms of pathogenesis: enterotoxogenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC) and diffusely adherent *E. coli* (DAEC).

Enterotoxogenic *E. coli* is a major cause of traveler's diarrhea, infantile diarrhea, and can cause diarrhea in young animals, although humans are the main reservoir in transmission to humans (35, 88). This organism colonizes the mucosal surface of the proximal small intestine through the use of a variety of colonization factors. At least 21 colonization factors have been characterized in human ETEC strains which include fimbrial and afimbrial adhesins (137). During and after colonization, ETEC strains produce heat-liable enterotoxin (LT) and/or heat-stable enterotoxin (ST). These toxins act upon separate intracellular targets, adenylate cyclase in the case of LT and guanylate cyclase in the case of ST, both leading to host cell secretion resulting in diarrhea (113).

Enteroaggregative *E. coli* are recognized for causing acute diarrhea in developing countries and persistent diarrhea worldwide (35, 67). During colonization of the intestine, these organisms demonstrate a stacked-brick like adherence that is mediated by an aggregative adherence fimbria (87, 88). The extent to which this fimbria plays in pathogenesis is not known as both pathogenic and non-pathogenic strains produce this structure. To date, it has not been demonstrated that EAEC produce ST or LT enterotoxins (68). EAEC produce several toxins such as the enteroaggregative heat stable toxin-1, a homologue of *E. coli* LT. Although many EAEC toxins have homologues, their exact mechanism of action, effect on the host, and role in pathogenesis have yet to be studied in depth. No one factor has been implicated in virulence of every pathogenic EAEC strain (68).

Enteroinvasive *E. coli* cause an infection that results in non-bloody, watery diarrhea and is similar to infections caused by *Shigella* spp. (35, 88). These organisms are related to *Shigella* spp. biochemically, genetically, and pathogenically (98). Enteroinvasive *E. coli* invade and proliferate in colonic epithelial cells with genes necessary for invasion encoded episomally on the plasmid pInv (88, 116). A type three secretion system (T3SS) encoded on this plasmid is responsible for translocation of effector proteins necessary for uptake and endocytic vacuole lysis. Enteroinvasive *E. coli* infection spreads via release of bacteria by apoptosis, lysis, and/or by migration from one epithelial cell to another via actin synthesis at one pole of a bacterial cell similar to *Listeria* and *Shigella* (88). Similar to ETEC, humans are the main reservoir of EIEC although a high infectious dose limits the risk of human-to-human transmission (57, 88).

Diffusely adherent *E. coli* are characterized by their diffuse adherence to Hep-2 cells, but association of DAEC with diarrhea disease is weak compared to other pathogenic *E. coli*. Little is known about DAEC mechanisms of pathogenesis or factors required for virulence (35, 88).

Enteropathogenic *E. coli* was the first virotype of *E. coli* to be described (68). These microorganisms were once an important cause of epidemic and sporadic infant diarrhea worldwide and are still frequently found in developing countries (88). The

hallmark of EPEC infection is the formation of attaching and effacing (A/E) lesions on the intestinal epithelium of humans and animals (2, 64, 71, 86, 96, 104, 118, 129, 139, 142). This intestinal histopathology is the result of a loss of the intestinal brush border, intimate adherence of the EPEC to the intestinal epithelial cell, and an actin rearrangement within the intestinal epithelial cells to form a pedestal (68, 88). There are many virulence factors and pathogenic mechanisms that are shared between EPEC and EHEC, which will be discussed later.

Enterohemorrhagic *E. coli* are a subset of the shiga-like toxin producing *E. coli* (STEC), which are also known as verotoxogenic *E. coli* (VTEC). These microorganisms are similar to EPEC in that they produce A/E lesions but elicit more severe symptoms in humans such as hemorrhagic colitis potentially leading to fatal sequelae. One of the most important serotypes of EHEC, *E. coli* O157:H7, is discussed below including descriptions of virulence factors common among all EHEC (68, 88).

Escherichia coli O157:H7

Escherichia coli O157:H7, an emerging foodborne pathogen, is a significant cause of foodborne illness in the United States. The Centers for Disease Control and Prevention estimates that there were 73,480 cases of illness attributed to *E. coli* O157:H7 in the United States, resulting in 2,186 hospitalizations and 61 deaths in 1997. The annual economic impact of *E. coli* O157:H7 is estimated at $405.2 million dollars (43).

The modern discovery of *E. coli* O157:H7 occurred in mid-1982 with two outbreaks linked to a national restaurant chain serving hamburgers, one in Oregon and a second in Michigan (23). Patients demonstrated a gastrointestinal illness characterized with severe abdominal cramps, low-grade fever and grossly bloody diarrhea later deemed hemorrhagic colitis. An organism cultured from the majority of 47 patients and from food common to the patients was identified as a rare serotype of *E. coli* called serotype O157:H7. Since this first documented outbreak of *E. coli* O157:H7, a number of outbreaks in human populations have occurred making this “rare” serotype an important cause of hemorrhagic colitis in the U.S. and other developed countries.

Escherichia coli O157:H7 human infection generally occurs as sporadic cases (1) rather than as outbreaks (35, 88). This is due to the spread of the organism via the fecal-oral route by food usually contaminated with ruminant feces. Many different types of food have been implicated as the source of outbreaks including undercooked hamburger, milk, apple juice, watermelons, green onions, and spinach (9, 51, 131). Outbreaks have also been associated with petting zoos, dairy farms, or recreation water sources near ruminant livestock yards (61, 62).

Ingestion of *E. coli* O157:H7 by humans is generally followed by a 3-4 day period while the organism colonizes and multiplies in the large bowel. The infectious dose for *E. coli* O157:H7 is very low, estimated at only 10-100 organisms (42). Symptoms are first manifested as abdominal cramping and non-bloody diarrhea with bloody stools developing 2-3 days later (93). The infection is usually self-limiting and resolves within a week after presentation of symptoms. Roughly 6% of cases develop hemolytic uremic syndrome (HUS), a dangerous sequelae that is currently the primary cause of acute kidney failure in children. HUS is a condition that is characterized by hemolytic anemia, thrombocytopenia, and uremia caused by shiga toxins produced by *E. coli* O157:H7 (131).

Reservoirs

Ruminants such as sheep and cattle are the major reservoirs of *E. coli* O157:H7, with cattle being the primary reservoir in the United States. There are also reports of *E. coli* O157:H7 being isolated from animals other than ruminants such as dogs, birds, and swine. In ruminants, *E. coli* O157:H7 can persist in experimentally infected animals up to two months as detected by fecal shedding (13, 25, 27-29, 31, 39, 121, 145, 146). However, this has been difficult to determine experimentally because of the intermittent nature of the shedding in infected animals. The lower gastrointestinal tract has been implicated as the site of *E. coli* O157:H7 colonization and persistence. Recent studies have linked long term colonization sites to the terminal rectal mucosa (89), but additional studies are needed to confirm this.

Research involving rumen carriage of *E. coli* O157:H7 has produced conflicting results. Early studies show that *E. coli* O157:H7 demonstrates unrestricted growth in rumen fluid of fasted cattle while growth is inhibited in the rumen fluid of well-fed cattle (99). This sparked research on the effect of the cattle diet on *E. coli* O157:H7 persistence; in several studies, it was isolated from the rumen of experimentally-infected cattle anywhere from 1 week (50) to 1 month (13).

In attempts to reduce or eliminate *E. coli* O157:H7 persistence in ruminants, many different abiotic and biotic factors have been studied. The effect of cattle diet on *E. coli* O157:H7 persistence and shedding is a contentious issue among researchers. The idea that volatile fatty acids produced during carbohydrate fermentation in the rumen inhibits *E. coli* O157:H7 growth was demonstrated *in vitro* (37), but *in vivo* studies have produced conflicting results (56, 134). The effect that pH has on *E. coli* O157:H7 survival in an artificial rumen model appears minimal (132). The effect of supplements such as monesin (143) and esculin (38) have had limited success in preventing *E. coli* O157:H7 carriage in the rumen but has yet to be proven in the field. Results from studies focused on biotic factors have suggested that indigenous rumen microflora inhibit growth or decrease survival of *E. coli* O157:H7 by competitive exclusion (20, 132). There is a significant gap in the understanding of the biology of *E. coli* O157:H7 in the rumen environment.

In summary, a greater understanding of the life physiology of *E. coli* O157:H7 in the rumen such as carbohydrate utilization, pH resistance, quorum sensing, protozoan carriage, or cross talk with other genera is needed. Such knowledge of the physiological state of *E. coli* O157:H7 before entering the lower gastrointestinal tract will be essential to a better understanding of the process of colonization and persistence in cattle populations.

Major virulence factors

The two virotypes EHEC and EPEC are distinct from other pathogenic *E. coli* in that they produce a unique pathology, called attaching and effacing (A/E) lesions when attached to intestinal epithelial cells (68, 88). The only other organism known to encode

genes for production of A/E lesions is the enterobacteria *Citrobacter rodentium.* This pathology appears as a loss of the intestinal brush border and microvilli along with the formation of a pedestal-like structure arising from the epithelial cell at the site of *E. coli* attachment. This interaction is thought to be a primary means of adherence in humans (68) and has been shown to facilitate colonization in adult ruminants (24), calves (29), and gnotobiotic piglets (33, 85, 138). A major pathogenicity island, known as the locus of enterocyte effacement or LEE, contains all genes known to play a role in the formation of A/E lesions. When transformed with LEE, *E. coli* K-12 and commensal *E. coli* display the A/E phenotype (84). LEE contains *eae*, which encodes the outer membrane protein intimin. This protein mediates adherence to intestinal epithelial cells in coordination with the translocated intimin receptor (Tir) (80). Tir is LEE encoded and can be transported by the LEE encoded type three secretion system (T3SS) (35, 67). Once translocated into an epithelial cell via the T3SS, Tir spans the host cell membrane allowing intimin to bind to its extracellular domain (30, 70, 80). Actin cytoskeletal rearrangement results from interaction of host proteins with the cytoplasmic domains of Tir (78). It has been shown that this association of Tir with intimin elicits an intracellular response that results in pedestal formation (66). It should be noted that Tir in EHEC is functionally distinct from Tir in EPEC as highlighted by work demonstrating that EPEC producing EHEC Tir did not elicit pedestal formation (18, 32, 69). It is now known that certain host factors are required for EPEC but not for EHEC Tir mediated actin polymerization (53) and that a LEE derived effector encoded on a separate bacteriophage is required for EHEC actin polymerization (17, 19). Intimin has been shown to mediate attachment in a Tir independent fashion by binding to the endogenous host cell receptor nucleolin with the same affinity but lower avidity when expressed on HEp-2 cells (114). Additionally, intimin has been demonstrated to bind ß1-integrins, although there is evidence to the contrary (77).

In addition to intimin, EHEC has few other purported afimbrial adhesins. A truncated form of the *efa-1* gene has been shown to play a role in adherence *in vitro* but not in calves or sheep (120). A complete copy homolog of *efa-1* encoded within pO157,

toxB, has demonstrated similar roles in adherence as *efa-1* (120, 126). Mutations in *efa-1* in EHEC O111:H$^-$ resulted in decreased binding to CHO cells and decreased colonization of the bovine intestine (91, 121). Clearly the role of these afimbrial adhesins in *E. coli* O157:H7 attachment and colonization needs further study.

Type three secretion systems have been proven to be virulence factors in many gram-negative plant and animal pathogens (35). These secretion systems are complex needle and syringe structures that are imbedded in the inner membrane, span the periplasmic space, are embedded in outer membrane and extend roughly 200 nm from the cell (47). The T3SS of *E. coli* O157:H7 is encoded within LEE and has been shown to be responsible for translocation of effector proteins across the bacterial membranes and the eukaryotic cell membrane into the host cell cytosol. The T3SS is comprised of approximately 20 proteins that are encoded within LEE (47). The basal body of the T3SS that is found within the inner membrane is comprised of proteins EscR, EscS, EscT, EscU, and EscV. The lipoprotein EscJ connects both the inner and outer membrane ring structures. The outer membrane ring is comprised of multiple EscC subunits. The needle structure is made of multiple subunits of EspA, which is anchored to the outer membrane ring via EscF. EscD and EscB are found on the distal end of the T3SS and are responsible for formation of the translocation pore in the host cell membrane (47). Seventeen proteins have been shown to be translocated via the T3SS thus they are called effector proteins. Tir, EspG, and TccP are translocated effectors that have been shown to play a role in pedestal formation while 14 of the proteins have yet to be identified functions (35).

Most EHEC encode for and produce one or two Shiga toxins. These toxins are so named due to their high homology with the toxins of *Shigella dysenteriae*. In EHEC they are encoded on either cryptic or functional prophage located within O-islands (35, 88). There are two distinct types of Shiga-like toxins that are categorized by being non-immunologically cross-reactive to one another; meaning antibody that is protective against one type of toxin is not protective against the other type (67). Major types of Shiga-like toxins in EHEC include Stx1, Stx2, and Stx2e although there are others such as Stx1c, Stx1d, Stx2c, Stx2d, Stx2f, and Stx2g (46). These toxins are AB subunit toxins

consisting of a single protoxin A subunit and 5 B subunits, the latter of which are responsible for toxin transport. The A subunit is not toxic until it is cleaved when transported from the Golgi apparatus into the host cell cytosol, which produces the active form of the toxin. Shiga toxins are N-glycosidases that, in the case of EHEC, cleave an adenine reside from the 28S ribosomal RNA bringing a halt to protein synthesis. These toxins are thought to elicit localized bleeding within the intestinal vasculature resulting in hemorrhagic colitis. Shiga-like toxins have been shown to find their way into the vascular system. The mechanism by which this occurs is currently being studied. Renal tissue, especially the glomeruli, are rich in globotriaosyleramide 3 (Gb_3,), the known receptor of Shiga-like toxins within humans (81). Shiga-like toxins are thought not to affect cattle as they lack Gb_3 in their vasculature (97). It has been recently shown that bovine crypt cells produce CD77 like molecules that may act as receptors for Stx1 although these cells are resistant to the apoptotic effects of the toxin (119). Swine are susceptible to a systemic disease caused by host adapted *E. coli* that produce the Stx2e variant of Shiga toxin and F18 fimbriae (22).

Escherichia coli O157:H7 has an outer membrane that is adorned with lipopolysaccharide (LPS) containing the O-antigen O157 (11, 21). LPS increases the cytotoxicity of *E. coli* O157:H7 *in vitro* as does the LPS of other bacteria, although the role in pathogenesis has yet to be fully explored (88). Interestingly it has been shown that the LPS of *E. coli* O157:H7 may interfere with attachment to epithelial cells as LPS mutants adhere better than wild type *in vitro* (11, 21). Due to the large number of potential fimbrial and afimbrial adhesins, LPS may only serve to interfere with attachment *in vivo*.

Profiling of the plasmids harbored within *E. coli* O157:H7 have shown that a 60-megadalton plasmid is present in 99-100% of strains assayed (76, 94, 100). Subsequent sequencing and analysis of the plasmid pO157 revealed a 92 kb F-like plasmid with 100 open reading frames (ORFs) (16). In addition to the potential adhesin *toxB* (discussed above), pO157 encodes a type two secretion system (110), enterohemolysin (111),

catalase-peroxidase (15), serine protease (14), lymphocyte inhibitory factor (63), and a metalloprotease (73).

Escherichia coli O157:H7 contain multiple gene clusters that have been predicted to encode fimbrial adhesions (58, 95), four of which are present in *E. coli* O157:H7 but not *E. coli* K-12 (79). There are two non-identical loci in *E. coli* O157:H7 that encode for long polar fimbriae (LPF) which are homologous to the LPF that are found in *Salmonella* spp. (137). When cloned into *E. coli* K-12, one of these loci confer slight adherence in cell culture and elaboration of LPF (135) while the other locus actually reduced adherence to cell culture (136). Mutations in one or both of the *lpf* loci result in reduced colonization in swine and sheep animal models (65), although it has been reported that these loci do not play a role in EPEC adherence (127).

A small number of *E. coli* O157:H7 strains produce thin aggregative fimbriae called curli under certain environmental conditions (140). The sequence, structure and regulation of curli seem to be highly conserved among *E. coli* and *Salmonella* (103). Expression of curli genes has been shown to be under the control of CsgD, RpoS, and a myriad of other regulators (3, 55, 92, 105) and increased expression of *csgD* leads to increased virulence in mice and increased invasion of HEp-2 cells by *E. coli* O157:H7 (141). Mutations in the curli operon demonstrated a slight decrease in adherence to HT-29 cells and showed no change in adherence to bovine intestinal tissue (105).

Clearly there are a number of virulence factors of *E. coli* O157:H7 that have been well characterized but still need further investigation. Commonalities between EPEC and EHEC virulence factors should be recognized but not accepted as truth as demonstrated by the mechanistic differences in pedestal formation between the two virotypes. More interesting are the large number of genes found within O-islands that have yet to be characterized.

Diagnostic and food safety considerations

Escherichia coli O157:H7 exhibits uncommon characteristics when compared to other *E. coli* strains. *Escherichia coli* O157:H7 does not grow well in nutrient broth over 44.5˚C, does not ferment sorbitol, is unable to produce β-glucuronidase, possess LEE,

and contains pO157. These characteristics are useful in the identification of *E. coli* O157:H7 (35).

Escherichia coli O157:H7 grows optimally in nutrient media at 37˚C and growth slows dramatically at 44 or 45˚C (36). This organism has been shown to survive in ground beef at -20˚C and -80˚C for up to 9 months (36). *Escherichia coli* O157:H7 has been isolated from surface water and can survive, especially in cold water, for several months (101). These cells are in a viable but non-culturable state as they do not grow on nutrient agar but can be seen to be intact by acridine orange staining (144).

Transcriptional analysis of Escherichia coli O157:H7 using microarrays

Transcriptome analysis of *E. coli* O157:H7 using microarrays has been limited with only a handful of studies in the past few years and even fewer using whole genome arrays. Most of the array-based studies of *E. coli* O157:H7 focus on rapid detection. The few transcriptional analysis studies reported have mainly focused on gene expression changes due to an environmental stressor such as high pressure (82) or during attachment of eukaryotic cell membranes (26). For instance, during colonizing of the intestinal tract of humans and ruminants, *E. coli* O157:H7 comes in close contact with host cell membranes. Microarray analysis of *E. coli* O157:H7 adhered to eukaryotic cell membranes revealed that stress responses are induced and that genes responsible for attachment and translation are repressed after 5 hours of association (26).

While in the intestinal tract of humans, *E. coli* O157:H7 is exposed to many different compounds including molecules that act as cellular signals. The *E. coli* Affymetrix 2.0 chips were used to investigate the effects of such molecules including epinephrine, norepinephrine, and indole on *E. coli* O157:H7 biofilms (5). Results indicated that epinephrine and norepinephrine up-regulated genes involved in virulence and surface colonization while indole down-regulated genes involved in these processes. Similarly, the effects of compounds derived from the oxidation of indole, namely 7-hydroxyindole and isatin, on the gene expression of *E. coli* O157:H7 during biofilm growth has been studied (74). During biofilm growth 20 genes in the locus of enterocyte effacement were repressed as well as genes previously shown to be involved in biofilm

growth and colonic acid biosynthesis. 7-hydroxyindole treatment results in the down-regulation of genes involved in purine nucleotide biosynthesis, cysteine biosynthesis, carbamoyl phosphate synthesis. Isatin repressed genes involved in AI-2 transport and indole biosynthesis while inducing flagellar genes and various transport genes.

Ultra high pressure is used in food processing as a method of preserving food by reducing or eliminating bacterial populations. To determine the effect of high pressure on bacterial gene expression, microarrays were used to compare normal and high pressure resistant *E. coli* O157:H7 strains at normal pressure to high pressure treatment (82). Genes involved in spontaneous mutation, thiol-disulfide redox, Fe-S cluster assembly, and stress responses were induced. No *E. coli* O157:H7 specific genes were represented on the array platform (*E. coli* K-12 Affymetrix gene chip) thus genes that may be involved with high pressure resistance and are encoded on O-islands or on pO157 would not be seen.

Although several characterizations of the global transcriptome of *E. coli* O157:H7 have been conducted more are needed. The global picture of transcript levels within *E. coli* O157:H7 during conditions *in vivo* either during colonization in a ruminant or during human infection may provide a wealth of knowledge regarding survival and pathogenesis in these environments. Besides uncovering new potential virulence factors, global regulators, and gene expression pathways, genomic tools may aid in the development of possible vaccines or identify new drug targets. These studies need to move towards *in vivo* like models or even within the *in vivo* environment, which may prove to be more relevant in the real world.

Heat shock

One common factor between all of these host environments is an elevation of temperature. This may be a shift from environmental ambient temperature to internal host temperature during the initial infection process or an increase in internal host temperature (pyrexia) due to bacteremia and inflammation. This pathogen is also subjected to extreme temperature conditions in the food processing industry during cleaning of equipment. Heat shock (HS) has the potential to be detrimental due to an increase in unfavorable

protein folding interactions. The HS response in *E. coli* K-12 is well studied (52) but the *E. coli* O157:H7 EDL933 genome contains additional genes in O-islands and plasmids (95), which respond to HS in unknown ways.

The HS response is often considered a general stress condition. This response has been widely studied in many organisms including *E. coli* K-12. Most, if not all, organisms elicit a response to HS by a rapid, selective, and transient synthesis of a number of universally conserved proteins known as HS proteins (6, 7). The HS response was first observed (102) and initially characterized (133) with work on the salivary glands of *Drosophila*. The HS response in *E. coli* was characterized independently in the same year (75, 147). Most characterized HS proteins function to degrade, fold, or re-fold proteins under normal and HS conditions (52). The alternative sigma factor (σ^{32}) encoded by *rpoH* has been shown to mediate the HS response in *E. coli*. Its role is to control the HS response by directing RNA polymerase to transcribe members of the HS regulon (75, 115, 127). Products of two major chaperone gene groups, *dnaK-dnaJ-grpE* and *groEL-ES*, are responsible for folding of nascent polypeptide chains, prevention of protein aggregation, and the re-folding of proteins that have lost their native confirmation due to thermal damage (52, 112). Heat shock causes an increase in unfavorable protein interactions requiring the need for such proteins. In addition to their functions in protein folding, these chaperone groups regulate σ^{32} by affecting its activity (45, 128). The HS protease HflB (often referred to as FtsH) also functions in a negative feedback loop by affecting the stability of σ^{32} (54, 60, 128). Other proteases such as Lon and members of the Clp family have not been shown to affect σ^{32} but do participate in the HS response (52). Small HS proteins IbpA and IbpB have been found to associate with inclusion bodies (59) and stabilize thermally aggregated proteins to promote disaggregation (83). HS proteins are not relegated to being expressed after an abrupt temperature increase as they are generally up-regulated as temperature increases. Effects on cellular responses such as HS by the unique information encoded within O-islands and on the plasmid pO157 are poorly understood.

Anaerobiasis

Intrinsic to most, if not all, enteric pathogens is the fact that they spend the majority of their time in an oxygen-limiting environment while inside the host. These organisms shift from a generally oxygen replete environment to one where oxygen is in short supply. While subjected to anaerobic or limiting oxygen environments *E. coli* has been shown to have differences in growth and metabolic rates as well as an altered transcriptome during anaerobiasis (106, 107). Microorganisms transitioning through these environments must alter transcriptional profiles in order to use simple sugars in anaerobic respiration and fermentation, and they must use alternate terminal electron receptors. Whether it is inside the human or ruminant, *E. coli* O157:H7 is not an exception to the rule as it resides within the gastrointestinal tract in oxygen limiting environments. The transcriptional response of *E. coli* K-12 (106, 107) and a small number *E. coli* O157:H7 genes (34) to anaerobiasis has been described, whereas the total transcriptome of *E. coli* O157:H7 under anaerobic conditions has yet to be studied.

Microarrays

History

Microarrays were developed in the late 1980's with experiments involving multi-analyte immunoassays by P. Ekins and coworkers (40, 41). Through a series of papers this group brought the idea of "reactions localized in 'microspots' distributed on an inert solid support". This idea was then carried through to the nucleic acid world in the mid 1990's with M. Schena and colleagues publishing the first DNA microarray paper with the differential expression of 45 genes in *Arabidopsis* using a two-color microarray (109) and then to the characterization of a entire transcriptome of *Saccharomyces cerevisiae* in 1997 (72). This idea of exploiting ligand binding, at least in the nucleic acid realm, started much earlier with a British biologist, E. M. Southern, in 1975 describing the detection of specific DNA fragments using DNA gel electrophoresis, a technique we now know today as a Southern blot (117). This technique involves transferring restriction digested DNA to a nitrocellulose membrane from a DNA electrophoresis gel. The

membrane and the bound DNA targets were then subjected to hybridization with radiolabeled DNA probes. This allows for detection of sequences of similar identity to be detected in the mixed DNA population on the membrane. Both this technology and microarrays rely on the ability of nucleic acid to interact and bind via hydrogen bond formation to sequences of similar identity.

Since the mid 1990's the increase in the number of reports of DNA microarrays has been astonishing. This phenomenon has been driven by a number of factors. The accumulation of genome or cDNA sequences along with plummeting costs and an increase in sequencing quality due to new technology has allowed for almost 800 complete bacterial and 50 archeal genome sequences to be published by early 2009. This, coupled with commercially available glass substrates, target preparation and hybridization kits, data acquisition systems, and data analysis software, has allowed many laboratories to study their organism of choice using DNA microarrays.

Types of DNA microarrays

There are several types of DNA microarrays including spotted cDNA, spotted DNA, spotted oligonucleotides, and *in situ* synthesized oligonucleotides, which can be categorized by the origin of their DNA and how the DNA is deposited onto the substrate. The earliest microarrays used cDNA sequence information due to the lack of appreciable amounts of genome sequence in the production of cDNA from mRNA in a reverse-transcription reaction. This is now an antiquated technology as more cost effective ways of probe generation have been developed. Spotted DNA microarrays, another outgoing technology developed as genome sequence became readily available, rely on the generation of open reading frame-specific DNA by polymerase chain reaction, which is subsequently purified and spotted on the microarray substrate. Advances in DNA synthesis allowed for the development of spotted oligonucleotide microarrays in which oligonucleotides representing the open reading frames of an entire genome can be synthesized and deposited onto a substrate. This was actually the precursor to *in situ* oligonucleotide microarrays that have been developed using photolithography, which is described below (108).

Target preparation methods and hybridization

All microarray gene expression studies start with isolated RNA and end with labeled cDNA hybridized to a microarray. RNA can be isolated using various techniques including phenol chloroform extraction, column isolation, ion exchange and aqueous phase isolation. One of the most important aspects of microarray research is the quality and purity of RNA being used. Degraded RNA results in very low cDNA yields, poor labeling, and subsequently unusable results (108). Quality can be assessed using methods such as RNA gel electrophoresis or an Agilent Bioanalyzer. With either technology rRNA peaks should be clearly distinguishable. In the case of bacteria, the ratio of density or area of 23S rRNA to 16S rRNA bands or peaks, depending on the method, should be approximately 2:1 (90). Impurities in RNA samples include carryover chemicals, DNA, and proteins from isolation that have the ability to inhibit the reverse transcription reaction and can lead to incorrect quantitation using a spectrophotometer.

The synthesis of DNA from RNA called reverse transcription (4, 130), cDNA generation, or cDNA synthesis for microarrays requires the incorporation of some type of labeled nucleotide. There are two general types of labeling: direct and indirect. Direct labeling involves the incorporation a nucleotide that is conjugated to a detectable molecule such as a fluorescent dye, electron dense material, etc. The most prevalent type of detectable molecule is a fluorescent dye such as a Cy or Alexa dye. In this type of reaction the reverse transcriptase must incorporate not only the nucleotides but nucleotides conjugated to a dye molecule, which in some cases are much larger that the nucleotide themselves. This leads to array bias since smaller dye groups have higher rates of dye incorporation than larger dye groups. This dye effect must be compensated for in the microarray experimental design and statistical analysis. Indirect labeling relies upon nucleotides containing reactive groups such as amino-allyl, amino-hexyl, or streptavidin modifications. After cDNA purification, cDNA samples are subjected to a dye coupling reaction where modified nucleotides interact with reactive dye groups. In the case of amino modifications, ester groups on dye molecules bind to the amino group resulting in fluorescently labeled cDNA. Biotin containing cDNA's are fluorescently labeled in

reactions with streptavidin conjugated dyes or antibodies. In both direct and indirect labeling, the amount of fluorescently labeled nucleotide or modified nucleotide in the reaction mixtures must be determined empirically. Specifically, the ratio of the unlabeled nucleotide to its modified counterpart must allow for cDNA synthesis to occur without over labeling. Complementary DNA can be labeled at such a high efficiency that dye quenching occurs.

Priming during the reverse transcription reaction can be accomplished using gene-specific, random, or poly-dT primers depending on the source of RNA. Gene-specific primers must be designed in such a way that all genes of interest are adequately primed using a mixture of many different oligonucleotides, which can be a computation intensive task. Poly-dT primers prime the synthesis of cDNA from RNA with poly-A tails. This is adequate for experiments involving eukaryotic organisms but inadequate for bacteria, which contain RNA that is not polyadenylated. Random primers do not require sophisticated design and can be used in most biological systems. Problems can arise, however, when working with genomes with high or low % G+C content, as their RNA's are inadequately primed (108). Also, the majority of the label occurs in structural ribosomal and transfer RNAs since they make up approximately 96% of the total RNA in the cell (90).

Reverse transcriptase, also known as RNA dependent DNA polymerase, is a critical part of cDNA synthesis. Since it's discovery in 1970 independently by H. Temin (130) and D. Baltimore (4), improvements have been made to reverse transcriptase to increase processivity and decrease RNase activity. This enzyme synthesizes single-stranded DNA from RNA producing a DNA:RNA hybrid, which after completion of cDNA synthesis for microarrays the RNA is hydrolyzed using a base or degraded using RNaseH ensuring that only one copy of DNA is made from one strand of RNA.

Hybridization can be considered the single most important step of the microarray technology. The technology depends on sequences of similar identity to bind one another and sequences that are dissimilar not binding to one another. This allows for the detection of a fluorescently labeled target bound to an unlabeled probe on the microarray substrate

surface after washing non-bound targets away. There are many factors that must be taken into consideration during hybridization including humidity, cover slip type, hybridization buffer composition, hybridization temperature, and labeled target concentration. Labeled targets resuspended in hybridization buffer are pipetted onto an array and covered with a cover slip. Slides are then placed in a hybridization chamber and placed in an incubator or water bath at a constant temperature for an optimized period of time. Slides are then washed with buffer containing salts, usually a solution of saline sodium citrate, along with a detergent such as sodium dodecyl sulfate. Washes are optimized for each type of microarray chemistry and type of DNA that is deposited on the slide. Washes generally start with less stringent washes with detergent leading to more and more stringent washes that do not contain a detergent. After washing, slides can be dried by centrifugation or by blowing particle free N_2 across the array surface. All of these steps must be determined empirically and optimized for the particular application.

Microarray bioinformatics

There are inherent multi-factorial and high dimensional problems associated with DNA microarray technology. The dawn of the computer age allowed for microarray technology to flourish by employing scripts and later commercially available software to aid in primer design, oligonucleotide design, data acquisition, data processing, statistical analysis, and database management.

Bioinformatics challenges start with PCR primer design for spotted PCR arrays and oligonucleotide design for spotted oligonucleotide arrays. For the generation of PCR products, primer pairs must bind to amplify a single region of the genome, should generate large yields of product of a preselected optimal size, and the product must be unique enough to ensure specific binding to only the portion of the ORF that was amplified. Oligonucleotides for oligonucleotide arrays are designed in much the same way except only one long oligonucleotide is usually needed for each ORF.

These bioinformatics challenges continue with scanning and image analysis. Once slides are dried post hybridization they must be scanned, as they are only useable for a couple days at most.

Data acquisition

Data acquisition relies upon commercially available microarray scanners that fall into two basic technologies based on how fluorescent dyes are excited and how emitted light is captured. The oldest and most ubiquitous microarray scanning technology uses lasers to excite fluorescent dyes incorporated into cDNA. During data acquisition a laser is passed across the microarray to excite fluorescently labeled cDNA bound to a spot. Light is emitted from the excited dye as the laser passes from one end of the slide to another, which is collected by a photomultiplier tube (PMT) and is quantified. Images are generated from these scans as 16 bit tiff files, which allow for any one pixel to be assigned intensity values of 0 to 2^{16}-1. For two color arrays, this process is repeated to gather data in a different dye channel using different lasers. The repertoire of fluorescent dyes that may be used is dictated by available lasers in the scanner and the range of excitation of the fluorophores. Different settings may be adjusted to optimize scanner performance and to provide the best possible images. Laser intensity may be varied to allow for more or less excitation of dyes depending on labeling efficiency, stringency of washes, and background present. Likewise the PMT gain may be varied to allow for increased or decreased detection of light excited from the slide. Scanning arrays multiple times with varying laser intensity and PMT gain settings, enhances discovery of differentially expressed genes (115).

A second technology that is markedly different from microarray laser scanners uses white light to excite all dyes present on the slide. In these systems a while light is passed across the microarray slide exciting all fluorescent dyes coupled to cDNA bound to spots. Emitted light travels through a filter specifically chosen to exclude all wavelengths except that of a specific dye. This light travels to a camera that photographs that specific region of the slide. Filters can be moved quickly in front of the camera to allow for collection of multiple dye channels on a specific region of the slide almost simultaneously. This process is repeated in increments across the slide using nano-motion technology. Individual images are stitched together *in silico* to form a single image of the slide. In this case, available filters and not laser excitation wavelength dictate the

repertoire of fluorescent dyes that can be imaged. Exposure times may be adjusted for each dye channel to compensate for differing labeling efficiencies, stringency of washes, and background.

Data processing and statistics

Once images are acquired, data processing can be carried out using spot finding software. To determine spot positions, a file produced from slide printing with assigned locations may be used or in some cases users may manually produce grids to be used in spot localization. A commonly used file format called a GAL file, or GenePix Array List, contains information about number and dimensions of subarrays, number and order of spots, and information regarding the genomic sequence that spot represents. The software is then able to determine spot location and carry out image segmentation. Image segmentation is the process by which pixels are defined as spot, background, or neither depending on settings used. Many different attributes of each spot on the array such as mean spot intensity and median background intensity can be extracted from the image and exported from the software to be used in statistical analysis. Spot finding software is increasingly being written to reduce manual processing with the capability of operating in a high throughput manner.

Microarrays pose numerous statistical questions and challenges. This is due to many factors including the type of array, the array application, experimental factors, and the inherently large amount of data that microarrays produce. In order to mitigate these challenges, planning before performing experiments is key. While designing an experiment, several factors must be taken into account including cost or availability of experimental units, cost of microarrays, plausible numbers of experimental units, types of treatment, etc. Increasing the number of biological replications should be favored over increasing technical replication in an experiment. The bias inherent in having two dyes that fluoresce at different intensities can be mitigated. Building dye swaps into the experimental design in order to have each treatment labeled the same number of times by each dye reduces this bias. This type of bias can also be reduced if a method such as indirect labeling is chosen instead of direct labeling.

Microarray data is not normally distributed and is subject to many different types of bias. Therefore it is necessary to reduce bias and normalize data to satisfy the assumptions of certain statistical tests such as the t-test. For the most part microarray data is log transformed to make the data more normal or to "normalize" the data. There are numerous sources of variation in experiments that we as biological researchers are not interested in. These sources include heat/light sensitivity of dyes, efficiency of dye incorporation, differences in amount of labeled cDNA hybridized to each channel of a slide, replicate slides, hybridization conditions, scanning conditions, etc. To reduce this type of non-biological variation, certain normalization methods are used. In this sense normalization is a misnomer in that we are not making the data more normal to fit assumptions of a particular statistical test. We are attempting to reduce the amount of non-biological variation to be able to better estimate the biological variation. There are many ways of normalizing the data including locally weighted polynomial regression (LOWESS) normalization (148), median centering, quantile normalization (12), and scale normalization (148). Although it is clear that normalization affects the final results, it is often not clear which normalization strategy is best.

Before statistical analysis it may be prudent to carry out background correction. Depending on microarray hybridization, washing, and scanning, slides can have differing amounts of non-specific fluorescence across the entire slide, portions of a slide, or fluorescent anomalies that are not part of a spot. These should be accounted for in some way and may be during spot localization by flagging a spot as abnormal and not used in subsequent statistical analysis. Background subtraction is a simple way for accounting for non-specific fluorescence and is done simply by subtracting the spot intensity value from the background intensity value generated during spot finding.

A statistical test must be carried out to determine significant differences in gene expression due to treatment. One test that is commonly carried is called a mixed model analysis. This includes factors that are fixed such as treatment effect, dye effect and factors that are random such as slide and interaction effects.

The probability of type I error increases when you increase the number of comparisons. The more statistical tests carried out, the greater the likelihood of finding a gene that is statistically significant when it is actually not. There are many methods for controlling type I error such as Bonferroni or Tukey (108), although controlling family-wise error rate in such a manner may be considered too conservative to apply to microarrays. In an ideal situation both type I and type II error should be reduced as much as possible. Benjamini and Hochberg in 1995 introduced the false discovery rate (FDR) as an alternative error rate that can be useful in microarray experiments (8). Experimenters can then choose what FDR they are willing to accept, which for publication of most microarray data is 10%. Many methods for estimating FDR have been published since (48, 122-125).

Validation of results

As with any experiment, using different methods to test the same hypothesis is good scientific practice. Microarray data should be validated using a means other than microarray. Generally, microarrays are validated using quantitative reverse transcription real time PCR (qRT-PCR), but it can also be done using older techniques such as a Northern blot. This involves selecting genes that are of interest and having both up- and down-regulated genes represented.

Quantitative reverse transcription real time PCR involves the production of cDNA from RNA using a reverse transcriptase followed by PCR using a method of detecting increasing amounts of DNA at the end of each cycle. This is the general paradigm of qRT-PCR although there are many different variations. This technique is more sensitive than others such as Northern blots or microarrays with the ability to detect a transcript in a sample that has been diluted 1:5,000,000 (44).

Quantitative reverse transcription real time PCR can be performed by either a one-step or two-step procedure. The one-step procedure involves having the reverse transcription reaction along with the PCR in a single reaction. This is usually done with incubation prior to PCR for the reverse transcription reaction followed by a heat step at or around 95˚C to inactivate the reverse transcriptase and activate the hot-start DNA

polymerase for PCR. Two-step qRT-PCR requires two sets of reactions that are completed separately and usually involves a purification step between the two reactions. Two-step reactions generally take more time and resources as the reactions are done in separate tubes.

There are several different ways of detecting increasing amounts of DNA during the PCR portion of qRT-PCR. There are two widely used chemistries that rely on fluorescence for detection. The cheapest and most widely used method is the integrating dye SYBR green. SYBR green is a nucleic acid stain commonly used in molecular biology that binds the minor groove of double stranded DNA at which time it becomes fluorescent. SYBR green is non-specific in that it will bind any dsDNA including contaminating genomic DNA or primer dimers generated during the PCR phase of the reaction. The SYBR green chemistry is relatively inexpensive, it works with any target gene, and it is sensitive because multiple dye molecules bind the target DNA strand.

The other main chemistry relies on an internal primer that contains a fluorophore and a quencher molecule. These are known either as molecular beacons or Taqman probes, each having specific design features. Molecular beacons exist in a hairpin conformation unless bound to a specific sequence generated during the PCR reaction. The hairpin structure positions the fluorophore and quencher in close proximity so that absorbed light energy does not result in fluorescence. As the concentration of the full-length product increases during the PCR reaction, the molecular beacon binds an internal sequence separating the fluorophore and quencher resulting in increasing fluorescence in direct relationship to the PCR product concentration. Taqman probes function differently in that the dual labeled oligonucleotide binds to the target sequence and does not contain a hairpin structure. It is also shorter in length allowing efficient transfer of energy from the fluorophore to the quencher in its normal conformation. During the extension reaction, the probe is hydrolyzed, effectively separating the fluorophore and quencher molecules and resulting in fluorescence.

These molecular probes have several advantages including increased specificity in the reaction since hybridization between the probe and the internal PCR product target

sequences must occur to result in fluorescence. It also allows multiplexing options with multiple probes possessing different dyes allowing for detection of multiple targets in one reaction. The single greatest disadvantage of molecular probes is that probes must be designed and synthesized for each target sequence investigated, increasing overall costs of the reaction. Probes are generally more expensive than primers used in SYBR green reactions, but in a diagnostic environment, these additional costs are not significant. They can be important, however, if multiple targets are under study and these targets change constantly such as in a research environment.

Whether or not using either one-step or two- step and SYBR green or molecular probes there must be a means to excite fluorescent dyes and detect that fluorescence while cycling through the varying temperatures required to carry out PCR. This requires the basic functions of a thermocycler coupled with a light and fluorescence detection system. Most real-time PCR machines are thermocyclers that have a light mounted above the block that excites fluorescent dyes over a large range of wavelengths. Fluorescence detection occurs by using a light detector mounted in the same position as the light with various filters to filter all but a single wavelength of light specific to emission wavelength for that fluorescent dye.

Summary with Goals and Hypothesis

Escherichia coli O157:H7 is generally in a category set apart from other foodborne pathogens. This is due to the severity of disease, possibility of fatal complications, and the small infectious dose. Although significant progress has been made over the last two decades in the understanding of its pathogenic mechanisms the need for more research still exists.

The overall hypothesis of this study is that *E. coli* O157:H7 increases its fitness for various compartments of the ruminant gastrointestinal tract by differential gene expression. Therefore, the goal of this study was to characterize changes in mRNA transcript levels using models of the bovine GI tract, heat shock and induction of stress related genes, anaerobiasis, exposure to rumen fluid under anaerobiasis, and internalization within protozoa (*Acanthamoeba*). Not all conditions could be adequately

modeled by *in vitro* conditions, however. Our studies confirmed previously identified genes in non-pathogenic *E. coli* strains under some conditions, but they also showed some striking differences and unexpected findings as shown in this dissertation.

References

1. 2007. Preliminary FoodNet data on the incidence of infection with pathogens transmitted commonly through food--10 states, 2006. MMWR Morb. Mortal. Wkly. Rep. **56:**336-339.
2. **Andrade, J. R., V. F. Da Veiga, M. R. De Santa Rosa, and I. Suassuna.** 1989. An endocytic process in HEp-2 cells induced by enteropathogenic *Escherichia coli*. J. Med. Microbiol. **28:**49-57.
3. **Arnqvist, A., A. Olsen, and S. Normark.** 1994. Sigma S-dependent growth-phase induction of the *csgBA* promoter in *Escherichia coli* can be achieved *in vivo* by sigma 70 in the absence of the nucleoid-associated protein H-NS. Mol. Microbiol. **13:**1021-1032.
4. **Baltimore, D.** 1970. RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature **226:**1209-1211.
5. **Bansal, T., D. Englert, J. Lee, M. Hegde, T. K. Wood, and A. Jayaraman.** 2007. Differential effects of epinephrine, norepinephrine, and indole on *Escherichia coli* O157:H7 chemotaxis, colonization, and gene expression. Infect. Immun. **75:**4597-4607.
6. **Bardwell, J. C. A., and E. A. Craig.** 1987. Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichia coli*. Proc. Natl. Acad. Sci. U S A **84:**5177-5181.
7. **Bardwell, J. C. A., and E. A. Craig.** 1984. Major heat shock gene of *Drosophila* and the *Escherichia coli* heat-inducible *dnaK* gene are homologous. Proc. Natl. Acad. Sci. U S A **81:**848-852.
8. **Benjamini, Y., and Y. Hochberg.** 1995. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. J. R. Statist. Soc. B **57:**289-300.

9. **Besser, R. E., S. M. Lett, J. T. Weber, M. P. Doyle, T. J. Barrett, J. G. Wells, and P. M. Griffin.** 1993. An outbreak of diarrhea and hemolytic uremic syndrome from *Escherichia coli* O157:H7 in fresh-pressed apple cider. JAMA **269:**2217-2220.

10. **Bettelheim, K. A.** 1986. Commemoration of the publication 100 years ago of the papers of Dr. Th. Escherich in which are described for the first time the organisms that bear his name. Zentralbl. Bakteriol. Mikrobiol. Hyg. **261:**255-265.

11. **Bilge, S. S., J. C. Vary, Jr., S. F. Dowell, and P. I. Tarr.** 1996. Role of the *Escherichia coli* O157:H7 O side chain in adherence and analysis of an *rfb* locus. Infect. Immun. **64:**4795-4801.

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

13. **Brown, C. A., B. G. Harmon, T. Zhao, and M. P. Doyle.** 1997. Experimental *Escherichia coli* O157:H7 carriage in calves. Appl. Environ. Microbiol. **63:**27-32.

14. **Brunder, W., H. Schmidt, and H. Karch.** 1997. EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. Mol. Microbiol. **24:**767-778.

15. **Brunder, W., H. Schmidt, and H. Karch.** 1996. KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic *Escherichia coli* O157:H7. Microbiology **142:**3305-3315.

16. **Burland, V., Y. Shao, N. T. Perna, G. Plunkett, H. J. Sofia, and F. R. Blattner.** 1998. The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. Nucleic Acids. Res. **26:**4196-4204.

17. **Campellone, K. G., M. J. Brady, J. G. Alamares, D. C. Rowe, B. M. Skehan, D. J. Tipper, and J. M. Leong.** 2006. Enterohaemorrhagic *Escherichia coli* Tir requires a C-terminal 12-residue peptide to initiate EspF-mediated actin assembly and harbours N-terminal sequences that influence pedestal length. Cell Microbiol. **8:**1488-1503.

18. **Campellone, K. G., A. Giese, D. J. Tipper, and J. M. Leong.** 2002. A tyrosine-phosphorylated 12-amino-acid sequence of enteropathogenic *Escherichia coli* Tir binds the host adaptor protein Nck and is required for Nck localization to actin pedestals. Mol. Microbiol. **43:**1227-1241.

19. **Campellone, K. G., D. Robbins, and J. M. Leong.** 2004. EspFU Is a Translocated EHEC Effector that Interacts with Tir and N-WASP and Promotes Nck-Independent Actin Assembly. Dev. Cell **7:**217-228.

20. **Chaucheyras-Durand, F., J. Madic, F. Doudin, and C. Martin.** 2006. Biotic and abiotic factors influencing in vitro growth of *Escherichia coli* O157:H7 in ruminant digestive contents. Appl. Environ. Microbiol. **72:**4136-4142.

21. **Cockerill, F., 3rd, G. Beebakhee, R. Soni, and P. Sherman.** 1996. Polysaccharide side chains are not required for attaching and effacing adhesion of *Escherichia coli* O157:H7. Infect. Immun. **64:**3196-3200.

22. **Cohen, A., G. E. Hannigan, B. R. Williams, and C. A. Lingwood.** 1987. Roles of globotriosyl- and galabiosylceramide in verotoxin binding and high affinity interferon receptor. J. Biol. Chem. **262:**17088-17091.

23. **Control, C. f. D.** 1982. Isolation of *E. Coli* O157:H7 from sporadic cases of hemorrhagic colitis.

24. **Cornick, N. A., S. L. Booher, and H. W. Moon.** 2002. Intimin facilitates colonization by *Escherichia coli* O157:H7 in adult ruminants. Infect. Immun. **70:**2704-2707.

25. **Cray, W. C., Jr., and H. W. Moon.** 1995. Experimental infection of calves and adult cattle with *Escherichia coli* O157:H7. Appl. Environ. Microbiol. **61:**1586-1590.

26. **Dahan, S., S. Knutton, R. K. Shaw, V. F. Crepin, G. Dougan, and G. Frankel.** 2004. Transcriptome of enterohemorrhagic *Escherichia coli* O157 adhering to eukaryotic plasma membranes. Infect. Immun. **72:**5452-5459.

27. **Dean-Nystrom, E. A., B. T. Bosworth, W. C. Cray, Jr., and H. W. Moon.** 1997. Pathogenicity of *Escherichia coli* O157:H7 in the intestines of neonatal calves. Infect. Immun. **65:**1842-1848.

28. **Dean-Nystrom, E. A., B. T. Bosworth, and H. W. Moon.** 1997. Pathogenesis of O157:H7 *Escherichia coli* infection in neonatal calves. Adv. Exp. Med. Biol. **412:**47-51.

29. **Dean-Nystrom, E. A., B. T. Bosworth, H. W. Moon, and A. D. O'Brien.** 1998. *Escherichia coli* O157:H7 requires intimin for enteropathogenicity in calves. Infect. Immun. **66:**4560-4563.

30. **Deibe, C., S. Kramer, T. Chakraborty, and F. Ebel.** 1998. EspE, a novel secreted protein of attaching and effacing bacteria, is directly translocated into infected host cells, where it appears as a tyrosine-phosphorylated 90 kDa protein. Mol. Microbiol. **28:**463-474.

31. **DeRisi, J. L., V. R. Iyer, and P. O. Brown.** 1997. Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale. Science **278:**680-686.

32. **DeVinney, R., J. L. Puente, A. Gauthier, D. Goosney, and B. B. Finlay.** 2001. Enterohaemorrhagic and enteropathogenic *Escherichia coli* use a different Tir-based mechanism for pedestal formation. Mol. Microbiol. **41:**1445-1458.

33. **Donnenberg, M. S., S. Tzipori, M. L. McKee, A. D. O'Brien, J. Alroy, and J. B. Kaper.** 1993. The role of the *eae* gene of enterohemorrhagic *Escherichia coli* in intimate attachment *in vitro* and in a porcine model. J. Clin. Invest. **92:**1418-1424.

34. **Dowd, S. E., and H. Ishizaki.** 2006. Microarray based comparison of two *Escherichia coli* O157:H7 lineages. BMC Microbiol. **6:**30.

35. **Doyle, M. P., and L. R. Beuchat.** 2007. Food Microbiology: Fundamentals and Frontiers, 3rd ed. ASM Press, Washington, D.C.

36. **Doyle, M. P., and J. L. Schoeni.** 1984. Survival and growth characteristics of *Escherichia coli* associated with hemorrhagic colitis. Appl. Environ. Microbiol. **48:**855-856.

37. **Duncan, S. H., H. J. Flint, and C. S. Stewart.** 1998. Inhibitory activity of gut bacteria against *Escherichia coli* O157 mediated by dietary plant metabolites. FEMS Microbiol. Lett. **164:**283-288.

38. **Duncan, S. H., E. C. Leitch, K. N. Stanley, A. J. Richardson, R. A. Laven, H. J. Flint, and C. S. Stewart.** 2004. Effects of esculin and esculetin on the survival of *Escherichia coli* O157 in human faecal slurries, continuous-flow simulations of the rumen and colon and in calves. Br. J. Nutr. **91:**749-755.

39. **Dziva, F., P. M. van Diemen, M. P. Stevens, A. J. Smith, and T. S. Wallis.** 2004. Identification of *Escherichia coli* O157:H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis. Microbiology **150:**3631-3645.

40. **Ekins, R., F. Chu, and J. Micallef.** 1989. High specific activity chemiluminescent and fluorescent markers: their potential application to high sensitivity and 'multi-analyte' immunoassays. J Biolumin Chemilumin **4:**59-78.

41. **Ekins, R. P.** 1989. Multi-analyte immunoassay. J. Pharm. Biomed. Anal. **7:**155-168.

42. **FDA** 2003, posting date. The Bad Bug Book. [http://www.cfsan.fda.gov/~mow/intro.html]

43. **Frenzen, P. D.** 2007. An Online Cost Calculator for Estimating the Economic Cost of Illness Due to Shiga Toxin-Producing *E. coli* (STEC) O157 Infections. *In* USDA (ed.), vol. 28.

44. **Gallup, J. M., and M. R. Ackermann.** 2006. Addressing fluorogenic real-time qPCR inhibition using the novel custom Excel file system 'Focusfield2-6GallupqPCRSet-upTool-001' to attain consistently high fidelity qPCR reactions. Biol. Proceed. Online **8:**87-152.

45. **Gamer, J., G. Multhaup, T. Tomoyasu, J. S. McCarty, S. Rudiger, H. J. Schonfeld, C. Schirra, H. Bujard, and B. Bukau.** 1996. A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the *Escherichia coli* heat shock transcription factor sigma32. EMBO J. **15:**607-617.

46. **Garcia-Aljaro, C., M. Muniesa, J. E. Blanco, M. Blanco, J. Blanco, J. Jofre, and A. R. Blanch.** 2005. Characterization of Shiga toxin-producing *Escherichia coli* isolated from aquatic environments. FEMS Microbiol. Lett. **246:**55-65.

47. **Garmendia, J., G. Frankel, and V. F. Crepin.** 2005. Enteropathogenic and enterohemorrhagic *Escherichia coli* infections: translocation, translocation, translocation. Infect. Immun. **73:**2573-2585.

48. **Genovese, C., and L. Wasserman.** 2002. Operating Characteristics and Extensions of the False Discovery Rate Procedure. J. R. Statist. Soc. B **64:**499-517.

49. **Gibson, G. R., and M. B. Roberfroid.** 1995. Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. J. Nutr. **125:**1401-1412.

50. **Grauke, L. J., I. T. Kudva, J. W. Yoon, C. W. Hunt, C. J. Williams, and C. J. Hovde.** 2002. Gastrointestinal tract location of *Escherichia coli* O157:H7 in ruminants. Appl. Environ. Microbiol. **68:**2269-2277.

51. **Griffin, P. M., and R. V. Tauxe.** 1991. The epidemiology of infections caused by *Escherichia coli* O157:H7, other enterohemorrhagic *E. coli*, and the associated hemolytic uremic syndrome. Epidemiol. Rev. **13:**60-98.

52. **Gross, C. A.** 1996. Function and regulation of the heat shock proteins. *In* A. Bock, R. Curtiss III, J. B. Kaper, F. C. Neidhardt, T. Nystrom, K. E. Rudd, and C. L. Squired (ed.), EcoSal- *Escherichia coli* and *Salmonella*: celluar and molecular biology. ASM Press, Washington, D.C.

53. **Gruenheid, S., R. DeVinney, F. Bladt, D. Goosney, S. Gelkop, G. D. Gish, T. Pawson, and B. B. Finlay.** 2001. Enteropathogenic *E. coli* Tir binds Nck to initiate actin pedestal formation in host cells. Nat. Cell. Biol. **3:**856-859.

54. **Guisbert, E., C. Herman, C. Z. Lu, and C. A. Gross.** 2004. A chaperone network controls the heat shock response in *E. coli*. Genes Dev. **18:**2812-2821.

55. **Hammar, M., A. Arnqvist, Z. Bian, A. Olsen, and S. Normark.** 1995. Expression of two *csg* operons is required for production of fibronectin- and

congo red-binding curli polymers in *Escherichia coli* K-12. Mol. Microbiol. **18:**661-670.

56. **Harmon, B. G., C. A. Brown, S. Tkalcic, P. O. Mueller, A. Parks, A. V. Jain, T. Zhao, and M. P. Doyle.** 1999. Fecal shedding and rumen growth of *Escherichia coli* O157:H7 in fasted calves. J. Food. Prot. **62:**574-579.

57. **Harris, J. R., J. Mariano, J. G. Wells, B. J. Payne, H. D. Donnell, and M. L. Cohen.** 1985. Person-to-person transmission in an outbreak of enteroinvasive *Escherichia coli*. Am. J. Epidemiol. **122:**245-252.

58. **Hayashi, T., K. Makino, M. Ohnishi, K. Kurokawa, K. Ishii, K. Yokoyama, C. G. Han, E. Ohtsubo, K. Nakayama, T. Murata, M. Tanaka, T. Tobe, T. Iida, H. Takami, T. Honda, C. Sasakawa, N. Ogasawara, T. Yasunaga, S. Kuhara, T. Shiba, M. Hattori, and H. Shinagawa.** 2001. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. DNA Res. **8:**11-22.

59. **Henry, M. D., S. D. Yancey, and S. R. Kushner.** 1992. Role of the heat shock response in stability of mRNA in *Escherichia coli* K-12. J. Bacteriol. **174:**743-748.

60. **Herman, C., D. Thevenet, R. D'Ari, and P. Bouloc.** 1995. Degradation of sigma 32, the heat shock regulator in *Escherichia coli*, is governed by HflB. Proc. Natl. Acad. Sci. U S A **92:**3516-3520.

61. **Heuvelink, A. E., C. van Heerwaarden, J. T. Zwartkruis-Nahuis, R. van Oosterom, K. Edink, Y. T. van Duynhoven, and E. de Boer.** 2002. *Escherichia coli* O157 infection associated with a petting zoo. Epidemiol. Infect. **129:**295-302.

62. **Hildebrand, J. M., H. C. Maguire, R. E. Holliman, and E. Kangesu.** 1996. An outbreak of *Escherichia coli* O157 infection linked to paddling pools. Commun. Dis. Rep. CDR Rev. **6:**33-36.

63. **Janka, A., M. Bielaszewska, U. Dobrindt, and H. Karch.** 2002. Identification and distribution of the enterohemorrhagic *Escherichia coli* factor for adherence

(*efa1*) gene in sorbitol-fermenting *Escherichia coli* O157:H-. Int. J. Med. Microbiol. **292:**207-214.

64. **Jerse, A. E., J. Yu, B. D. Tall, and J. B. Kaper.** 1990. A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. Proc. Natl. Acad. Sci. USA **87:**7839-7843.
65. **Jordan, D. M., N. Cornick, A. G. Torres, E. A. Dean-Nystrom, J. B. Kaper, and H. W. Moon.** 2004. Long polar fimbriae contribute to colonization by *Escherichia coli* O157:H7 *in vivo*. Infect. Immun. **72:**6168-6171.
66. **Kalman, D., O. D. Weiner, D. L. Goosney, J. W. Sedat, B. B. Finlay, A. Abo, and J. M. Bishop.** 1999. Enteropathogenic *E. coli* acts through WASP and Arp2/3 complex to form actin pedestals. Nat. Cell Biol. **1:**389-391.
67. **Kaper, J. B.** 1998. Enterohemorrhagic *Escherichia coli*. Cur. Opin. Microbiol. **1:**103-108.
68. **Kaper, J. B., J. P. Nataro, and H. L. Mobley.** 2004. Pathogenic *Escherichia coli*. Nat. Rev. Microbiol. **2:**123-140.
69. **Kenny, B.** 2001. The enterohaemorrhagic *Escherichia coli* serotype O157:H7 Tir molecule is not functionally interchangeable for its enteropathogenic *E. coli* serotype O127:H6 homologue. Cell. Microbiol. **3:**499-510.
70. **Kenny, B., R. DeVinney, M. Stein, D. J. Reinscheid, E. A. Frey, and B. B. Finlay.** 1997. Enteropathogenic *E. coli* (EPEC) Transfers Its Receptor for Intimate Adherence into Mammalian Cells. Cell **91:**511-520.
71. **Knutton, S., M. M. Baldini, J. B. Kaper, and A. S. McNeish.** 1987. Role of plasmid-encoded adherence factors in adhesion of enteropathogenic *Escherichia coli* to HEp-2 cells. Infect. Immun. **55:**78-85.
72. **Lashkari, D. A., J. L. DeRisi, J. H. McCusker, A. F. Namath, C. Gentile, S. Y. Hwang, P. O. Brown, and R. W. Davis.** 1997. Yeast microarrays for genome wide parallel genetic and gene expression analysis. Proc. Natl. Acad. Sci. USA **94:**13057-13062.

73. **Lathem, W. W., T. E. Grys, S. E. Witowski, A. G. Torres, J. B. Kaper, P. I. Tarr, and R. A. Welch.** 2002. StcE, a metalloprotease secreted by *Escherichia coli* O157:H7, specifically cleaves C1 esterase inhibitor. Mol. Microbiol. **45:**277-288.

74. **Lee, J., T. Bansal, A. Jayaraman, W. E. Bentley, and T. K. Wood.** 2007. Enterohemorrhagic *Escherichia coli* Biofilms Are Inhibited by 7-Hydroxyindole and Stimulated by Isatin. Appl. Environ. Microbiol. **73:**4100-4109.

75. **Lemaux, P. G., S. L. Herendeen, P. L. Bloch, and F. C. Neidhardt.** 1978. Transient rates of synthesis of individual polypeptides in *E. coli* following temperature shifts. Cell **13:**427-434.

76. **Levine, M. M., J. G. Xu, J. B. Kaper, H. Lior, V. Prado, B. Tall, J. Nataro, H. Karch, and K. Wachsmuth.** 1987. A DNA probe to identify enterohemorrhagic *Escherichia coli* of O157:H7 and other serotypes that cause hemorrhagic colitis and hemolytic uremic syndrome. J. Infect. Dis. **156:**175-182.

77. **Liu, H., L. Magoun, and J. M. Leong.** 1999. Beta 1-Chain Integrins Are Not Essential for Intimin-Mediated Host Cell Attachment and Enteropathogenic *Escherichia coli*-Induced Actin Condensation. Infect. Immun. **67:**2045-2049.

78. **Liu, H., P. Radhakrishnan, L. Magoun, M. Prabu, K. G. Campellone, P. Savage, F. He, C. A. Schiffer, and J. M. Leong.** 2002. Point mutants of EHEC intimin that diminish Tir recognition and actin pedestal formation highlight a putative Tir binding pocket. Mol. Microbiol. **45:**1557-1573.

79. **Low, A. S., N. Holden, T. Rosser, A. J. Roe, C. Constantinidou, J. L. Hobman, D. G. Smith, J. C. Low, and D. L. Gally.** 2006. Analysis of fimbrial gene clusters and their expression in enterohaemorrhagic *Escherichia coli* O157:H7. Environ. Microbiol. **8:**1033-1047.

80. **Luo, Y., E. A. Frey, R. A. Pfuetzner, A. L. Creagh, D. G. Knoechel, C. A. Haynes, B. B. Finlay, and N. C. J. Strynadka.** 2000. Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. Nature **405:**1073-1077.

81. **Mainil, J.** 1999. Shiga/verocytotoxins and Shiga/verotoxigenic *Escherichia coli* in animals. Vet. Res. **30:**235-257.

82. **Malone, A. S., Y. K. Chung, and A. E. Yousef.** 2006. Genes of *Escherichia coli* O157:H7 that are involved in high-pressure resistance. Appl. Environ. Microbiol. **72:**2661-2671.

83. **Matuszewska, M., D. Kuczynska-Wisnik, E. Laskowska, and K. Liberek.** 2005. The small heat shock protein IbpA of *Escherichia coli* cooperates with IbpB in stabilization of thermally aggregated proteins in a disaggregation competent state. J. Biol. Chem. **280:**12292-12298.

84. **McDaniel, T. K., and J. B. Kaper.** 1997. A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. Mol. Microbiol. **23:**399-407.

85. **McKee, M. L., A. R. Melton-Celsa, R. A. Moxley, D. H. Francis, and A. D. O'Brien.** 1995. Enterohemorrhagic *Escherichia coli* O157:H7 requires intimin to colonize the gnotobiotic pig intestine and to adhere to HEp-2 cells. Infect. Immun. **63:**3739-3744.

86. **Moon, H. W., S. C. Whipp, R. A. Argenzio, M. M. Levine, and R. A. Giannella.** 1983. Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. Infect. Immun. **41:**1340-1351.

87. **Nataro, J. P., Y. Deng, D. R. Maneval, A. L. German, W. C. Martin, and M. M. Levine.** 1992. Aggregative adherence fimbriae I of enteroaggregative *Escherichia coli* mediate adherence to HEp-2 cells and hemagglutination of human erythrocytes. Infect. Immun. **60:**2297-2304.

88. **Nataro, J. P., and J. B. Kaper.** 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11:**142-201.

89. **Naylor, S. W., J. C. Low, T. E. Besser, A. Mahajan, G. J. Gunn, M. C. Pearce, I. J. McKendrick, D. G. Smith, and D. L. Gally.** 2003. Lymphoid follicle-dense mucosa at the terminal rectum is the principal site of colonization of

enterohemorrhagic *Escherichia coli* O157:H7 in the bovine host. Infect. Immun. **71:**1505-1512.

90. **Neidhardt, F. C., J. L. Ingraham, and M. Schaechter.** 1990. Physiology of the Bacterial Cell: A Molecular Approach. Sinauer Associates, Inc, Sunderland, Mass.

91. **Nicholls, L., T. H. Grant, and R. M. Robins-Browne.** 2000. Identification of a novel genetic locus that is required for *in vitro* adhesion of a clinical isolate of enterohaemorrhagic *Escherichia coli* to epithelial cells. Mol. Microbiol. **35:**275-288.

92. **Olsen, A., A. Arnqvist, M. Hammar, and S. Normark.** 1993. Environmental regulation of curli production in *Escherichia coli*. Infect. Agents Dis. **2:**272-274.

93. **Ostroff, S. M., P. M. Griffin, R. V. Tauxe, L. D. Shipman, K. D. Greene, J. G. Wells, J. H. Lewis, P. A. Blake, and J. M. Kobayashi.** 1990. A statewide outbreak of *Escherichia coli* O157:H7 infections in Washington State. Am. J. Epidemiol. **132:**239-247.

94. **Ostroff, S. M., P. I. Tarr, M. A. Neill, J. H. Lewis, N. Hargrett-Bean, and J. M. Kobayashi.** 1989. Toxin genotypes and plasmid profiles as determinants of systemic sequelae in *Escherichia coli* O157:H7 infections. J. Infect. Dis. **160:**994-998.

95. **Perna, N. T., G. Plunkett, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner.** 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature **409:**529-533.

96. **Polotsky, Y. E., E. M. Dragunskaya, V. G. Seliverstova, T. A. Avdeeva, M. G. Chakhutinskaya, I. Ketyi, A. Vertenyl, B. Ralovich, L. Emody, I. Malovics, N. V. Safonova, E. S. Snigirevskaya, and E. I. Karyagina.** 1977. Pathogenic

effect of enterotoxigenic *Escherichia coli* and *Escherichia coli* causing infantile diarrhoea. Acta. Microbiol. Acad. Sci. Hung. **24:**221-236.

97. **Pruimboom-Brees, I. M., T. W. Morgan, M. R. Ackermann, E. D. Nystrom, J. E. Samuel, N. A. Cornick, and H. W. Moon.** 2000. Cattle lack vascular receptors for *Escherichia coli* O157:H7 Shiga toxins. Proc. Natl. Acad. Sci. USA **97:**10325-10329.

98. **Pupo, G. M., R. Lan, and P. R. Reeves.** 2000. Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. Proc. Natl. Acad. Sci. USA **97:**10567-10572.

99. **Rasmussen, M. A., W. C. Cray, Jr., T. A. Casey, and S. C. Whipp.** 1993. Rumen contents as a reservoir of enterohemorrhagic *Escherichia coli*. FEMS Microbiol. Lett. **114:**79-84.

100. **Ratnam, S., S. B. March, R. Ahmed, G. S. Bezanson, and S. Kasatiya.** 1988. Characterization of *Escherichia coli* serotype O157:H7. J. Clin. Microbiol. **26:**2006-2012.

101. **Rice, E. W., C. H. Johnson, D. K. Wild, and D. J. Reasoner.** 1992. Survival of *Escherichia coli* O157:H7 in drinking water associated with a waterborne disease outbreak of hemorrhagic colitis. Lett. Appl. Microbiol. **15:**38-40.

102. **Ritossa, F.** 1963. New puffs induced by temperature shock, DNP and salicylate in salivary chromosomes of *Drosophila melanogaster*. Drosoph. Inf. Serv. **37:**122-123.

103. **Romling, U., Z. Bian, M. Hammar, W. D. Sierralta, and S. Normark.** 1998. Curli Fibers Are Highly Conserved between *Salmonella typhimurium* and *Escherichia coli* with Respect to Operon Structure and Regulation. J. Bacteriol. **180:**722-731.

104. **Rothbaum, R., A. J. McAdams, R. Giannella, and J. C. Partin.** 1982. A clinicopathologic study of enterocyte-adherent *Escherichia coli*: a cause of protracted diarrhea in infants. Gastroenterol. **83:**441-454.

105. **Saldaoa, Z., X. Xicohtencatl-Cortes, R. Avelino, A. D. Phillips, J. B. Kaper, J. L. Puente, and J. A. Girun.** 2009. Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing *Escherichia coli* and identification of Fis as a negative regulator of curli. Environ. Microbiol. **11:**992-1006.

106. **Salmon, K., S. P. Hung, K. Mekjian, P. Baldi, G. W. Hatfield, and R. P. Gunsalus.** 2003. Global gene expression profiling in *Escherichia coli* K12. The effects of oxygen availability and FNR. J. Biol. Chem. **278:**29837-29855.

107. **Salmon, K. A., S. P. Hung, N. R. Steffen, R. Krupp, P. Baldi, G. W. Hatfield, and R. P. Gunsalus.** 2005. Global gene expression profiling in *Escherichia coli* K12: effects of oxygen availability and ArcA. J. Biol. Chem. **280:**15084-15096.

108. **Schena, M.** 2003. Microarray Analysis. John Wiley & Sons, Inc., Hoboken, New Jersey.

109. **Schena, M., D. Shalon, R. W. Davis, and P. O. Brown.** 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science **270:**467-470.

110. **Schmidt, H., B. Henkel, and H. Karch.** 1997. A gene cluster closely related to type II secretion pathway operons of gram-negative bacteria is located on the large plasmid of enterohemorrhagic *Escherichia coli* O157 strains. FEMS Microbiol. Lett. **148:**265-272.

111. **Schmidt, H., H. Karch, and L. Beutin.** 1994. The large-sized plasmids of enterohemorrhagic *Escherichia coli* O157 strains encode hemolysins which are presumably members of the *E. coli* alpha-hemolysin family. FEMS Microbiol. Lett. **117:**189-196.

112. **Schroder, H., T. Langer, F. U. Hartl, and B. Bukau.** 1993. DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. EMBO J. **12:**4137-4144.

113. **Sears, C. L., and J. B. Kaper.** 1996. Enteric bacterial toxins: mechanisms of action and linkage to intestinal secretion. Microbiol. Rev. **60:**167-215.

114. **Sinclair, J. F., and A. D. O'Brien.** 2002. Cell Surface-localized Nucleolin Is a Eukaryotic Receptor for the Adhesin Intimin-gamma of Enterohemorrhagic *Escherichia coli* O157:H7. J. Biol. Chem. **277:**2876-2885.

115. **Skibbe, D. S., X. Wang, X. Zhao, L. A. Borsuk, D. Nettleton, and P. S. Schnable.** 2006. Scanning microarrays at multiple intensities enhances discovery of differentially expressed genes. Bioinformatics **22:**1863-1870.

116. **Small, P. L., and S. Falkow.** 1988. Identification of regions on a 230-kilobase plasmid from enteroinvasive *Escherichia coli* that are required for entry into HEp-2 cells. Infect. Immun. **56:**225-229.

117. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503-517.

118. **Staley, T. E., E. W. Jones, and L. D. Corley.** 1969. Attachment and penetration of *Escherichia coli* into intestinal epithelium of the ileum in newborn pigs. Am. J. Pathol. **56:**371-392.

119. **Stamm, I., M. Mohr, P. S. Bridger, E. Schropfer, M. Konig, W. C. Stoffregen, E. A. Dean-Nystrom, G. Baljer, and C. Menge.** 2008. Epithelial and Mesenchymal Cells in the Bovine Colonic Mucosa Differ in Their Responsiveness to *Escherichia coli* Shiga Toxin 1. Infect. Immun. **76:**5381-5391.

120. **Stevens, M. P., A. J. Roe, I. Vlisidou, P. M. van Diemen, R. M. La Ragione, A. Best, M. J. Woodward, D. L. Gally, and T. S. Wallis.** 2004. Mutation of *toxB* and a Truncated Version of the *efa-1* Gene in *Escherichia coli* O157:H7 Influences the Expression and Secretion of Locus of Enterocyte Effacement-Encoded Proteins but not Intestinal Colonization in Calves or Sheep. Infect. Immun. **72:**5402-5411.

121. **Stevens, M. P., P. M. van Diemen, G. Frankel, A. D. Phillips, and T. S. Wallis.** 2002. *efa1* Influences Colonization of the Bovine Intestine by Shiga Toxin-Producing *Escherichia coli* Serotypes O5 and O111. Infect. Immun. **70:**5158-5166.

122. **Storey, J. D.** 2002. A Direct Approach to False Discovery Rates. J. R. Statist. Soc. B **64:**479-498.

123. **Storey, J. D.** 2003. The Positive False Discovery Rate: A Bayesian Interpretation and the q-Value. Ann. Stat. **31:**2013-2035.

124. **Storey, J. D., and R. Tibshirani.** 2001. Estimating false discovery rates under dependence, with applications to DNA microarrays. Stanford University.

125. **Storey, J. D., and R. Tibshirani.** 2003. Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. USA **100:**9440-9445.

126. **Tatsuno, I., M. Horie, H. Abe, T. Miki, K. Makino, H. Shinagawa, H. Taguchi, S. Kamiya, T. Hayashi, and C. Sasakawa.** 2001. *toxB* Gene on pO157 of Enterohemorrhagic *Escherichia coli* O157:H7 is Required for Full Epithelial Cell Adherence Phenotype. Infect. Immun. **69:**6660-6669.

127. **Tatsuno, I., R. Mundy, G. Frankel, Y. Chong, A. D. Phillips, A. G. Torres, and J. B. Kaper.** 2006. The *lpf* gene cluster for long polar fimbriae is not involved in adherence of enteropathogenic *Escherichia coli* or virulence of *Citrobacter rodentium*. Infect. Immun. **74:**265-272.

128. **Tatsuta, T., T. Tomoyasu, B. Bukau, M. Kitagawa, H. Mori, K. Karata, and T. Ogura.** 1998. Heat shock regulation in the *ftsH* null mutant of *Escherichia coli:* dissection of stability and activity control mechanisms of sigma32 *in vivo.* Mol. Microbiol. **30:**583-593.

129. **Taylor, C. J., A. Hart, R. M. Batt, C. McDougall, and L. McLean.** 1986. Ultrastructural and biochemical changes in human jejunal mucosa associated with enteropathogenic *Escherichia coli* (0111) infection. J. Pediatr. Gastroenterol. Nutr. **5:**70-73.

130. **Temin, H. M., and S. Mizutani.** 1970. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature **226:**1211-1213.

131. **Thorpe, C. M.** 2004. Shiga toxin-producing *Escherichia coli* infection. Clin. Infect. Dis. **38:**1298-1303.

132. **Thran, B. H., H. S. Hussein, D. Redelman, and G. C. Fernandez.** 2003. Influence of pH treatments on survival of *Escherichia coli* O157:H7 in continuous cultures of rumen contents. Exp. Biol. Med. (Maywood) **228:**365-369.

133. **Tissieres, A., H. Mitchell, and U. Tracy.** 1974. Protein synthesis in salivary glands of *Drosophila melanogaster*: relation to chromosome puffs. J. Mol. Biol. **84:**389-398.

134. **Tkalcic, S., C. A. Brown, B. G. Harmon, A. V. Jain, E. P. Mueller, A. Parks, K. L. Jacobsen, S. A. Martin, T. Zhao, and M. P. Doyle.** 2000. Effects of diet on rumen proliferation and fecal shedding of *Escherichia coil* O157:H7 in calves. J. Food. Prot. **63:**1630-1636.

135. **Torres, A. G., J. A. Giron, N. T. Perna, V. Burland, F. R. Blattner, F. Avelino-Flores, and J. B. Kaper.** 2002. Identification and Characterization of *lpfABCC'DE*, a Fimbrial Operon of Enterohemorrhagic *Escherichia coli* O157:H7. Infect. Immun. **70:**5416-5427.

136. **Torres, A. G., K. J. Kanack, C. B. Tutt, V. Popov, and J. B. Kaper.** 2004. Characterization of the second long polar (LP) fimbriae of *Escherichia coli* O157:H7 and distribution of LP fimbriae in other pathogenic *E. coli* strains. FEMS Microbiol. Lett. **238:**333-344.

137. **Torres, A. G., X. Zhou, and J. B. Kaper.** 2005. Adherence of Diarrheagenic *Escherichia coli* Strains to Epithelial Cells. Infect. Immun. **73:**18-29.

138. **Tzipori, S., F. Gunzer, M. S. Donnenberg, L. de Montigny, J. B. Kaper, and A. Donohue-Rolfe.** 1995. The role of the *eaeA* gene in diarrhea and neurological complications in a gnotobiotic piglet model of enterohemorrhagic *Escherichia coli* infection. Infect. Immun. **63:**3621-3627.

139. **Tzipori, S., H. Karch, K. I. Wachsmuth, R. M. Robins-Browne, A. D. O'Brien, H. Lior, M. L. Cohen, J. Smithers, and M. M. Levine.** 1987. Role of a 60-megadalton plasmid and Shiga-like toxins in the pathogenesis of infection caused by enterohemorrhagic *Escherichia coli* O157:H7 in gnotobiotic piglets. Infect. Immun. **55:**3117-3125.

140. **Uhlich, G. A., J. E. Keen, and R. O. Elder.** 2001. Mutations in the *csgD* Promoter Associated with Variations in Curli Expression in Certain Strains of *Escherichia coli* O157:H7. Appl. Environ. Microbiol. **67:**2367-2370.

141. **Uhlich, G. A., J. E. Keen, and R. O. Elder.** 2002. Variations in the *csgD* Promoter of *Escherichia coli* O157:H7 Associated with Increased Virulence in Mice and Increased Invasion of HEp-2 Cells. Infect. Immun. **70:**395-399.

142. **Ulshen, M. H., and J. L. Rollo.** 1980. Pathogenesis of *Escherichia coli* gastroenteritis in man--another mechanism. N. Engl. J. Med. **302:**99-101.

143. **Van Baale, M. J., J. M. Sargeant, D. P. Gnad, B. M. DeBey, K. F. Lechtenberg, and T. G. Nagaraja.** 2004. Effect of forage or grain diets with or without monensin on ruminal persistence and fecal *Escherichia coli* O157:H7 in cattle. Appl. Environ. Microbiol. **70:**5336-5342.

144. **Wang, G., and M. P. Doyle.** 1998. Survival of enterohemorrhagic *Escherichia coli* O157:H7 in water. J. Food. Prot. **61:**662-667.

145. **Woodward, M. J., A. Best, K. A. Sprigings, G. R. Pearson, A. M. Skuse, A. Wales, C. M. Hayes, J. M. Roe, J. C. Low, and R. M. La Ragione.** 2003. Non-toxigenic *Escherichia coli* O157:H7 strain NCTC12900 causes attaching-effacing lesions and *eae*-dependent persistence in weaned sheep. Int. J. Med. Microbiol. **293:**299-308.

146. **Woodward, M. J., D. Gavier-Widen, I. M. McLaren, C. Wray, M. Sozmen, and G. R. Pearson.** 1999. Infection of gnotobiotic calves with *Escherichia coli* O157:H7 strain A84. Vet. Rec. **144:**466-470.

147. **Yamamori, T., K. Ito, Y. Nakamura, and T. Yura.** 1978. Transient regulation of protein synthesis in *Escherichia coli* upon shift-up of growth temperature. J. Bacteriol. **134:**1133-1140.

148. **Yang, Y. H., S. Dudoit, P. Luu, D. M. Lin, V. Peng, J. Ngai, and T. P. Speed.** 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucl. Acids. Res. **30:**e15.

CHAPTER 2. TRANSCRIPTOME ANALYSIS OF *ESCHERICHIA COLI* O157:H7 DURING HEAT SHOCK

A paper published in FEMS Microbiology Letters*

Michael D. Carruthers and F. Chris Minion

Department of Veterinary Microbiology and Preventive Medicine, and Interdepartmental Microbiology Program, Iowa State University, Ames, IA 50011

Abstract

The response to elevated temperature (heat shock) is highly conserved. The transcriptome of *Escherichia coli* K-12 has been studied over a variety of conditions while such studies involving *E. coli* O157:H7 are just now being conducted. To better understand the impact of heat shock on *E. coli* O157:H7, global transcript levels of strain EDL933 cells shifted from 37°C to 50°C for 15 min were compared to cells left at 37°C by microarray. Using a mixed model analysis, 193 genes were found to be differentially transcribed at $p < 0.0042$ with a q-value of < 0.1. The 111 down-regulated genes include the curli pili associated genes *csgABCDEFG,* maltose transport associated proteins *malEFK,* and NADH dehydrogenase subunit encoding *nuoCEHIJN*. The 82 genes up-regulated include the heat shock induced genes *rpoH*, *dnaK*, *dnaJ*, *groEL*, *groES*, and *grpE* along with two LEE encoded genes; hypothetical gene Z5121 and *sepZ*. Twenty three additional genes located in O-islands were found to be differentially expressed. qRT-PCR was performed to validate microarray results. Also, samples subjected to a 30˚C to 42˚C shift were examined by qRT-PCR to confirm differential transcription of selected genes. These results indicate that this pathogen may regulate its virulence factors in response to temperature changes.

---

* Reprinted with permission of FEMS Microbiol. Lett. 2009, 295, 96-102.

Introduction

Since its isolation and identification over two decades ago (Riley*, et al.*, 1983), *Escherichia coli* O157:H7 has remained a significant foodborne pathogen. It causes acute gastrointestinal disease in humans which manifests from mild diarrhea to hemorrhagic colitis with the possibility of developing hemolytic uremic syndrome, a potentially fatal sequelae that is the leading cause of acute renal failure in children (Nataro & Kaper, 1998). Upon entering a host, *E. coli* O157:H7 is subjected to a variety of environments that must be endured to ultimately colonize and/or cause disease.

One common factor between all of these host environments is an elevation of temperature. This may be a shift from environmental ambient temperature to internal host temperature during the initial infection process or an increase in internal host temperature (pyrexia) due to bacteremia and inflammation. This pathogen is also subjected to extreme temperature conditions in the food processing industry. Heat shock (HS) has the potential to be detrimental due to an increase in unfavorable protein folding interactions. The HS response in *E. coli* K-12 is well studied (Gross, 1996), but the *E. coli* O157:H7 EDL933 genome contains additional genes in O-islands and plasmids (Perna*, et al.*, 2001), which respond to HS in unknown ways.

The HS response is often considered a general stress condition. It has been widely studied in many organisms including *E. coli* K-12. Most, if not all, organisms elicit a response to HS by a rapid, selective, and transient synthesis of a number of universally conserved proteins known as HS proteins (Bardwell & Craig, 1984, Bardwell & Craig, 1987). Effects on cellular responses such as HS by the unique information encoded within O-islands and on the plasmid pO157 are poorly understood.

The focus of this study was to characterize the transcriptional response of *E. coli* O157:H7 to HS. To accomplish this, the transcriptional profile of pathogenic *E. coli* O157:H7 EDL933 during heat-shock was compared to cells grown under normal conditions using two-color microarrays and quantitative real time PCR (qRT-PCR). The results of this experiment demonstrate that while classical HS proteins were up-regulated as expected, genes not previously identified as being differentially expressed during HS

and genes not found within *E. coli* K-12 were differentially expressed in *E. coli* O157:H7 during HS as well.

Materials and Methods

Strains and culture conditions

Escherichia coli O157:H7 EDL933 was used in this study. To obtain cultures for study, 100 ml of Luria-Bertani broth in 250 ml Erlenmeyer flasks were inoculated with 1 ml of an overnight culture. After 3 h incubation at 37°C and shaking at 200 rpm, half of the flasks were shifted from 37°C to 50°C for 15 min. One milliliter of cells were quickly pelleted by centrifugation at 12,000 × *g* for 45 sec, 500 µl of culture supernate was aspirated, 1 ml of RNAprotect Bacteria Reagent (Qiagen, Valencia, CA) was added, pellets were homogenized by vortexing, samples were incubated at room temperature for 10 min, and then RNA was isolated. An additional study was performed with a temperature upshift from 30°C to 42°C, which was analyzed by semi-quantitative real-time polymerase chain reaction (qRT-PCR) for specific genes.

Microarray construction

The *E. coli* O157:H7 microarray consists of 85.7 % of the opening reading frames (ORFs) encoded within the *chromosome* of *E. coli* O157:H7 EDL933 (5453 ORFs) and pO157 (101 ORFs) represented as 100-500 bp PCR products (probes) spotted to glass substrates. Determination of primer sequences and generation and purification of PCR products were carried out as previously described (Madsen*, et al.*, 2006). Primers were purchased from Integrated DNA Technologies (Coralville, IA). A custom printer at the Microarray Facility at Washington University was used to print the microarrays onto Corning UltraGAPS substrates (Corning, Big Flats, NY). The array contains 48 sub-arrays in a 15-column by 15-row configuration with probes spotted in duplicate in a non-contiguous fashion. Slides were UV cross-linked at 450 mJ in a Stratalinker (Stratagene, La Jolla, CA) to immobilize the DNA. A description of the array can be found at the

National Center for Biotechnology Information Gene Expression Omnibus (Edgar*, et al.*, 2002) under platform GPL6272.

RNA isolation

RNA was isolated from samples using the RNeasy Mini Kit (Qiagen) following the lysis and digestion kit protocol. Samples were treated using of Turbo DNase (Ambion, Austin, TX) as per manufacturers instructions. The lysis and digestion protocol was followed with two 50 µl RNase-free water elutions. Each sample was treated with 2 µl of DNase (Ambion, Austin, TX) at 37°C for 30 min. Samples were purified and concentrated using Microcon YM-30 columns (Millipore, Billerica, MA), and the quantity and purity were determined using a ND-1000 spectrophotometer (Nanodrop, Wilmington, DE). Samples were determined to be free of contaminating genomic DNA by the absence of a band on a DNA electrophoresis gel after 30 rounds of PCR as follows. Samples were diluted to 100 ng μl^{-1} and used as template in PCR with native Taq DNA polymerase (Invitrogen, Carlsbad, CA) as per manufacturer's instructions. Primers used were forward primer 5'-CAGGGTTCTGAGCAGGTGTG-3' and reverse primer 5'-GACAACCCGGACTGGGAG-3'. Thermocycler conditions were 94°C for 5 min, then 39 cycles of 94°C for 30 sec, 62°C for 30 sec, 72°C for 30 sec, with a final 72°C for 5 min. RNA sample integrity was assessed using the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) with all samples having a RNA integrity number of 9.0 and higher.

Labeled cDNA generation and hybridization

Targets were prepared as previously described excluding the specific hexamer primer set (Oneal*, et al.*, 2008). To prime the cDNA synthesis reactions, 10 µg of random hexamer (Integrated DNA Technologies) was used. Frequency of dye incorporation and yield was determined using the ND-1000 spectrophotometer. From each paired sample, 1.5 µg of labeled cDNA was combined, dried-down, resuspended in 60 µl Pronto! cDNA/long oligonucleotide hybridization solution (Corning), incubated at 95°C for 5 min and centrifuged at 13,000 × *g* for 2 min at room temperature. This

solution was then pipetted on a freshly pre-hybridized microarray slide, covered with a 22 x 60 mm HybriSlip (Grace Bio Labs, Bend, OR), placed in a Corning hybridization chamber, and incubated in a 42°C water bath for 12 to 16 h. Array pre-hybridization was conducted using sodium borohydride as previously described (Raghavachari*, et al.*, 2003). Post-hybridization washes were conducted according to Corning's UltraGAPS protocol, dried by centrifugation, and then scanned.

Microarray experimental design

Twelve heat-shocked RNA samples were paired with twelve control RNA samples for hybridization on a two-color microarray slide. Cy3 and Cy5 dye assignments to control and treated samples were reversed for six of the arrays to account for variation in labeling efficiencies (dye bias).

Image acquisition and data analysis

Scanning, image segmentation, and normalization were conducted as previously described (Madsen*, et al.*, 2006). The normalized values for duplicate spots were averaged within each array to produce one normalized measure of expression of fluorescence for each probe sequence and each of the 24 experimental units. Mixed model analysis was conducted as previously described (Madsen, et al., 2006), excluding slide region effects and slide-by-region interactions.

To discover potential regulatory motifs in upstream regions of significant differentially transcribed genes, a one block motif analysis was carried out using BioProspector and MEME in conjunction with BioOptimizer (Bailey & Elkan, 1994, Liu*, et al.*, 2001, Jensen & Liu, 2004). We limited the analysis to genes up-regulated that are not found in *E. coli* K-12 and have greater than 50 bp of upstream intergenic sequence (Z0968, Z2323, Z2400, Z2565, Z3341, hopD, sepZ, Z0402, Z1453, Z1456, Z1648, Z4321, Z4650, Z4883, Z5121). Motifs were deemed significant if identified by BioOptimizer using both MEME and BioProspector output as input. BPROM (www.softberry.com) was used for promoter prediction.

Validation of microarray data

To confirm significant transcriptional differences between genes in both temperature upshift samples, qRT-PCR was performed on nine genes (*csgB, dnaK, groEL, malF, rpoH, sepZ, yceP,* Z1443 and Z5121) shown to have significant transcriptional differences during heat shock as determined by microarrays. In addition to these genes, *clpB* was also analyzed by qRT-PCR with both temperature upshift samples. The Stratagene (La Jolla, CA) Brilliant II One Step qRT-PCR kit along with the Stratagene Mx3005P QPCR System was used for qRT-PCR. The ISU method was used to calculate fold change between treatment and control samples (Gallup & Ackermann, 2006). Table 2.1 describes the qRT-PCR primers used in this study.

Table 2.1. qRT-PCR Primers

Gene	Sequence	Tm °C
clpB-F	GGT GCG CGT TCT TAA TCT TTG CGA	60.2
clpB-R	CCT CCA CGC ATT TGT TCA ATC GCT	60.2
csgB-F	AGG TAG TAG CAA CCG GGC AAA GAT	60.9
csgB-R	GCA CCT TGC GAA ATA CTG GCA TCA	60.1
dnaK-F	TGG TGG TCA GAC TCG TAT GCC AAT	60.1
dnaK-R	ATT GCT ACA GCT TCG TCC GGG TTA	60.3
groEL-F	GTG GGT ATC AAA GTT GCA CTG CGT	60.0
groEL-R	TTT GGT TGG GTC CAG GAT ACC CAT	60.2
malF-F	ATT CAA CTG TTA ACC AAC GGC GGC	60.4
malF-R	TGG CTT TCA GGT TCA CTA TCG CCA	60.5
phoP-F	AGC TGG CAG GAC AAA GTC GAA GTA	60.2
phoP-R	CAT TCG CGC CAT CAC CTC TTC AAT	59.9
rpoH-F	TCG TCA AAG TTG CGA CCA CCA AAG	60.0
rpoH-R	ATC GTC GTC GGA AGA CAG GTC AAA	60.0
sepZ-F	TGG CGA CCT CAC TCA GTG GAA ATA	59.9
sepZ-R	CGG CTA TAA CTC TAA CGG TGC GAT	58.5
Z1443-F	AAA GCG CGA GGA AGT AAG CAA G	58.0
Z1443-R	TGT CAT CAG AAG GGC TTA TGA ACT	56.0
Z5121-F	CGT ACG CAG GGA GTG ATT GAA CAT	59.0
Z5121-R	CAT CCT GCG AAC GCG CTC AAT AAT	60.0
yceP-F	AGT CAT TCA GAC TCA TCC GCT CGT	59.8
yceP-R	TGG TAG TGC AAA CGC AAC ATC AGC	60.3

Results

Array development and construction

Of the 5,555 open reading frames encoded within the *E. coli* O157:H7 EDL933 genome and pO157, 4,756 were able to have unique PCR primers designed and be successfully amplified. Agarose gel electrophoresis was conducted post amplification and purification to determine specificity of primer pairs and product sizes. Arrays were quality controlled post printing using Spot QC (Integrated DNA Technologies). The hybridized arrays were scanned and spot morphology, spacing and overall array quality examined.

Heat Shock

To determine the effect of heat shock on the transcriptome of *E. coli* O157:H7 EDL933, microarrays were used to compare transcript levels within planktonic cultures subjected to elevated temperature to cultures held at a constant temperature. Cultures were at the same phase of growth. Extraction of total RNA from approximately 6.0×10^8 cells yielded a minimum of 25 µg of total RNA post-DNase treatment. Equal amounts of paired control and heat shock total RNA from all samples were subjected to cDNA synthesis and dye coupling to allow for an equal number of dye swaps. Data from each of twelve replicates were used in the statistical analysis.

Statistical analysis indicated that 193 genes demonstrated transcriptional differences with a p-value < 0.0042 and an estimated false discovery rate (FDR) < 10%. Significance and differences in transcript levels for all genes are depicted as a volcano plot (Fig. 2.1). Of the 193 genes differentially expressed, 82 genes were up-regulated while 111 genes were down-regulated. The 50 most significantly differentially expressed genes, 25 up- and 25 down-regulated are shown (Tables 2.2 and 2.3, respectively). The remaining genes are shown in supplemental Table 2.S1 (Appendix A). Table 2.4 lists genes that are differentially expressed that are not present in *E. coli* K-12. Nine genes, 7 up-regulated and 2 down-regulated, were chosen for validation using qRT-PCR along with *phoP*, a control gene that showed no differential expression. We also chose to

examine *clpB* by qRT-PCR since it had previously been shown to be up-regulated during heat shock (Gross, 1996). In all cases the qRT-PCR results confirmed the direction of the transcript difference and indicated that the microarrays show comparatively lower fold change values (Tables 2.2 and 2.3). In the case of *clpB*, the microarray failed to identify any changes, but its up-regulation was confirmed by qRT-PCR (Table 2.2). The dataset from the microarrays can be accessed from the Gene Expression Omnibus (NCBI) using series accession number GSE11463 (http://www.ncbi.nlm.nih.gov/geo/).

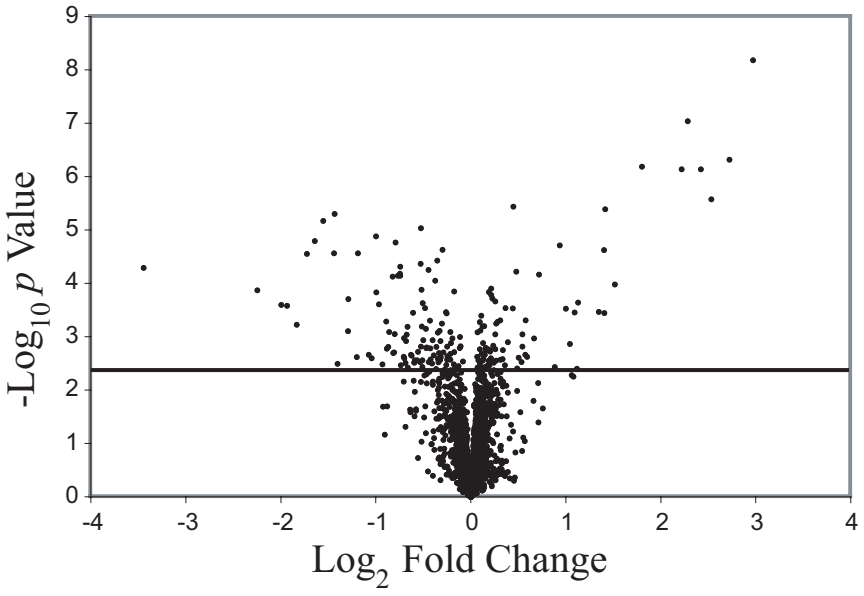

Figure 2.1. Volcano plot of transcriptional differences in *E. coli* O157:H7 EDL933 during heat shock. Individual differences are plotted as $\log_2$ Fold Change vs. $-\log_{10} p$ value. Points above the line at $p < 0.0042$ indicate differential expression at a FDR of 10%.

Table 2.2. Top twenty-five up-regulated genes based on *p* value of microarray data

ID	Locus	Product/Function	*p* value	*q* value	FC[1]	FC[2]	FC[3]
clpB	Z3886	Protein disaggregation chaperone				5.7	35.4
dinI	Z1698	Putative damage induced protein I	< 0.0004	0.002	1.4		
dnaJ	Z0015	Chaperone with DnaK heat shock protein	< 0.0004	0.001	4.7		
dnaK	Z0014	Molecular chaperone DnaK	< 0.0004	0.001	5.4	20.3	18.3
groEL	Z5748	Chaperonin GroEL	< 0.0004	0.001	6.6	12.8	13.7
groES	Z5747	Co-chaperonin GroES	< 0.0004	0.000	7.9		
grpE	Z3907	Heat shock protein GrpE	< 0.0004	0.002	2.7		
hflB	Z4540	Integral membrane peptidase	< 0.0004	0.022	2.2		
htpG	Z0590	Heat shock protein 90	< 0.0004	0.005	1.9		
htpX	Z2876	Heat shock protein HtpX	< 0.0004	0.000	4.9		
hydN	Z4021	Electron transport from formate to hydrogen Fe-S centers	< 0.0004	0.018	1.2		
ibpA	Z5183	Heat shock protein	< 0.0004	0.002	5.8		
ibpB	Z5182	Heat shock protein	< 0.0004	0.006	2.6		
prlC	Z4898	Oligopeptidase A	< 0.0004	0.025	1.3		
rpoH	Z4835	RNA polymerase sigma factor	< 0.0004	0.027	2.5	5.2	3.1
rrmJ	Z4541	Cell division protein	< 0.0004	0.014	2.9		
secG	Z4537	Protein-export membrane protein	< 0.0004	0.025	2.0		
slpA	Z0033	Probable peptidyl-prolyl cis-trans isomerase	< 0.0004	0.022	1.2		
topA	Z2536	DNA topoisomerase I	< 0.0004	0.010	1.4		
ybaP	Z0601	Putative ligase	< 0.0004	0.016	1.2		
yceP	Z1697	Hypothetical protein	< 0.0004	0.001	3.5	11.7	4.7
ygiM	Z4408	Hypothetical protein	< 0.0004	0.016	1.1		
Z1443[4]	Z1443	Unknown protein encoded by prophage BP-933W	< 0.0004	0.025	1.4		6.9
Z4321	Z4321	Putative PagC-like membrane protein	< 0.0004	0.020	1.2		
Z5745	Z5745	Hypothetical protein	< 0.0004	0.010	1.6		

[1] FC = Fold Change Microarray (37-50˚C)
[2] FC = Fold Change qRT-PCR (37-50°C)
[3] FC = Fold Change qRT-PCR (30-42°C)
[4] indicates not found in *E. coli* K-12.

Table 2.3. Top twenty-five down-regulated genes based on *p* value of microarray data

ID	Locus	Product/Function	*p* value	*q* value	FC[1]	FC[2]	FC[3]
aspA	Z5744	Aspartate ammonia-lyase	< 0.0002	0.002	-2.7		
crp	Z4718	Cyclic AMP receptor protein	< 0.0002	0.008	-1.4		
csgA	Z1676	Curlin major subunit	< 0.0002	0.006	-2.7		
csgB	Z1675	Minor curlin subunit	< 0.0002	0.010	-1.7	-13.4	-4.1
csgC	Z1677	Putative curli production protein	< 0.0002	0.010	-1.4		
cydB	Z0901	Cytochrome d terminal oxidase subunit II	< 0.0002	0.005	-1.7		
dctA	Z4942	C4-dicarboxylate transport protein	< 0.0002	0.016	-1.4		
fba	Z4263	Fructose-bisphosphate aldolase	< 0.0002	0.004	-2.0		
hdeA	Z4922	Hypothetical protein	< 0.0002	0.016	-4.7		
malF	Z5631	Maltose permease protein	< 0.0002	0.009	-10.9	-4.6	-20.3
ompR	Z4760	Osmolarity response regulator	< 0.0002	0.008	-1.3		
pepD	Z0298	Aminoacyl-histidine dipeptidase	< 0.0002	0.012	-1.3		
tdcB	Z4469	Threonine dehydratase	< 0.0002	0.005	-3.1		
tdH	Z5043	L-threonine 3-dehydrogenase	< 0.0002	0.010	-1.8		
yciD	Z2034	Putative outer membrane protein	< 0.0002	0.009	-1.7		
ydaA	Z2435	Hypothetical protein	< 0.0002	0.006	-2.3		
ydiA	Z2732	Hypothetical protein	< 0.0002	0.006	-1.2		
yeiA	Z3402	Dihydropyrimidine dehydrogenase	< 0.0002	0.010	-1.7		
Z0968[4]	Z0968	Unknown protein encoded by prophage CP-933K	< 0.0002	0.003	-2.9		
Z1479[4]	Z1479	Unknown protein encoded by prophage BP-933W	< 0.0002	0.016	-1.1		
Z2220	Z2220	Putative sensor kinase	< 0.0002	0.010	-1.7		
Z2565[4]	Z2565	Putative chaperone protein	< 0.0002	0.016	-2.0		
Z2869	Z2869	Hypothetical protein	< 0.0002	0.003	-1.4		
Z3332[4]	Z3332	Unknown protein encoded by prophage CP-933V	< 0.0002	0.006	-3.3		

[1] FC = Fold Change Microarray (37-50˚C)
[2] FC = Fold Change qRT-PCR (37-50°C)
[3] FC = Fold Change qRT-PCR (30-42°C)
[4] indicates not found in *E. coli* K-12

Table 2.4. Differentially transcribed genes not found in *E. coli* K-12

ID	Locus	Product/Function	[1]FC	[2]FC	*p* value	*q* value
Down-regulated						
Z0968	Z0968	Unknown protein encoded by prophage CP-933K	-2.9		0.000	0.003
Z1479	Z1479	Unknown protein encoded by prophage BP-933W	-1.1		0.000	0.016
Z1495	Z1495	Unknown protein encoded by prophage BP-933W	-1.1		0.001	0.052
Z2323	Z2323	Hypothetical protein	-1.3		0.001	0.057
Z2400	Z2400	Partial putative repressor protein encoded by prophage CP-933R	-1.3		0.003	0.079
Z2565	Z2565	Putative chaperone protein	-2.0		0.000	0.016
Z3332	Z3332	Unknown protein encoded by prophage CP-933V	-3.3		0.000	0.006
Z3341	Z3341	Unknown protein encoded by prophage CP-933V	-1.6		0.001	0.039
Z4333	Z4333	Putative cytotoxin	-1.1		0.002	0.067
Up-regulated						
hopD	Z4693	Putative leader peptidase	1.1		0.004	0.091
sepZ	Z5122	Type three secretion system associated protein	1.1	24.0	0.004	0.097
Z0402	Z0402	Putative beta-barrel outer membrane protein	1.2		0.001	0.066
Z1059	Z1059	Hypothetical protein	1.6		0.001	0.053
Z1338	Z1338	Putative DNA replication factor encoded by cryptic prophage CP-933M	1.1		0.004	0.091
Z1340	Z1340	Unknown protein encoded by cryptic prophage CP-933M	1.1		0.003	0.079
Z1443	Z1443	Unknown protein encoded by prophage BP-933W	1.4	6.9	0.000	0.025
Z1453	Z1453	Unknown protein encoded by prophage BP-933W	1.3		0.002	0.068
Z1456	Z1456	Unknown protein encoded by prophage BP-933W	1.4		0.004	0.092
Z1648	Z1648	Unknown in putative ISEc8	1.2		0.001	0.048
Z2394	Z2394	Unknown protein encoded by prophage CP-933R	1.2		0.003	0.079
Z4321	Z4321	Putative PagC-like membrane protein	1.2		0.000	0.020
Z4650	Z4650	Putative transferase	1.2		0.001	0.035
Z4883	Z4883	Hypothetical protein	1.2		0.002	0.067
Z5121	Z5121	Hypothetical protein	1.1	21.3	0.002	0.067

[1]FC = Fold Change Microarray (37-50˚C)
[2]FC = Fold Change qRT-PCR (30-42˚C)

In a second series of experiments, a heat shock study using a temperature shift from 30°C to 42°C was performed, and select genes were examined by qRT-PCR. In accordance to the 37°C to 50°C shift, all of the genes examined (*csgB, dnaK, groEL, malF, rpoH, sepZ, yceP*, Z1443 and Z5121) showed the same transcriptional shift (Tables 2 and 3). Additionally, *clpB* was not differentially transcribed on our microarrays but it was by qRT-PCR (Table 2.2).

No significant one block motifs were identified in the *E. coli* O157:H7 specific up-regulated genes that have greater than 50 bp of upstream intergenic sequence. All of these genes have the predicted sigma 70 promoter two block sequences upstream identified using BPROM.

Discussion

Overview

The data in this study clearly demonstrate that *E. coli* O157:H7 exhibits transcriptional differences in response to HS. Not surprisingly, previously identified *E. coli* K-12 HS genes were found to be up-regulated in O157:H7 as in other systems. Interestingly, genes common to *E. coli* O157:H7 and *E. coli* K-12 that had not previously been reported as differentially regulated in HS were identified. This may be attributed to different experimental conditions or array platforms, or to a real change in gene expression due to the difference in genomic composition. Further studies will be needed to differentiate these two possibilities. Some genes found within O-islands were also found to be differentially expressed as well.

When comparing our HS results with other microarray studies conducted on similar species, the fold change across all genes was slightly lower. A comparison of fold change across microarray platforms is difficult as there are many factors that can affect these differences. Our qRT-PCR results indicate that microarrays underestimate the fold change of transcripts, which is common (Morey*, et al.*, 2006).

Classic HS response

Many previously identified HS response genes were up-regulated during HS treatment. These include genes encoding major chaperone subunits *groEL, groES*, *dnaK, dnaJ, grpE,* sigma factors *rpoH, rpoD,* heat shock proteins *htpGX, ibpAB,* and proteases *clpX, ftsH, hslV, hflX* among others. These results indicate that the core HS response in *E. coli* O157:H7 is similar to that of *E. coli* K-12. This was expected as HS results in the increase of unfavorable protein interactions such as misfolding and aggregation. This requires a quick and sustained response to avoid detrimental effects by producing proteins involved in aiding protein folding, protein disaggregation, and proteolysis. Several previously identified HS genes were not identified in these studies, however, variation in growth conditions, treatment conditions, and microarray platforms may be potential reasons for this.

***E. coli* O157:H7 specific genes**

The sequencing of the *E. coli* K-12 and *E. coli* O157:H7 genomes revealed a 4.1 megabase pair (Mbp) shared backbone with *E. coli* O157:H7 EDL933 containing an additional 1.34 Mbp in O-islands with *E. coli* K-12 containing an additional 0.53 Mbp in K-islands (Blattner*, et al.*, 1997, Perna*, et al.*, 2001). Of the 1,379 genes encoded within O-islands 23 genes were differentially expressed, 15 up- and 9 genes down-regulated. Twelve of these genes are encoded on bacteriophage including 5 genes found within the Stx2 encoding bacteriophage BP-933W and 2 genes are within the Stx1 encoding bacteriophage CP-933V. No increase in transcripts related to the shiga-like toxins was observed, however. This indicates that the bacteriophage within *E. coli* O157:H7 are responding to HS, but this increase in temperature does not result in increased levels of shiga-toxin transcripts.

*Z*5121 and *sepZ*, genes adjacent to one another within the locus of enterocyte effacement, were found to be up-regulated during HS. Although it has no assigned function, the *sepZ* (*espZ*) gene product has been shown to be an effector protein transported via a type III secretion system into host cells by enteropathogenic *E. coli* (Kanack, et al., 2005). This is preliminary evidence that genes located in an important

pathogenesis island are up-regulated as a result of an increase in temperature. This could offer an advantage to a pathogen by only producing virulence factors when in the host and limit energy expenditures.

It is interesting to note that all genes assayed for regulatory sequences have predicted σ^{70} consensus sequences. The gene encoding σ^{70}, *rpoD*, was significantly up-regulated as a result of HS.

In conclusion, we have demonstrated that *E. coli* O157:H7 shows a similar gene expression pattern to *E. coli* K-12 in response to HS. Additionally, *E. coli* O157:H7 not only up-regulates genes found within O-islands but also up-regulates genes in the enterocyte effacement locus in response to HS. It is unlikely that these latter genes are involved in the classic view of the HS response (i.e., limiting unfavorable protein interactions), but rather are up-regulated during HS for another purpose such as pathogenesis. This will require further study.

Acknowledgements

FCM was funded for the microarray construction through United States Department of Agriculture Specific Cooperative Agreement 58-3625-2-127. We thank Melissa L. Madsen for microarray construction and advice during this project, and Jack Gallup for assistance with qRT-PCR.

References

Bailey TL & Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**: 28-36.

Bardwell JCA & Craig EA (1984) Major heat shock gene of *Drosophila* and the *Escherichia coli* heat-inducible *dnaK* gene are homologous. *Proc. Natl. Acad. Sci. U S A* **81**: 848-852.

Bardwell JCA & Craig EA (1987) Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A* **84**: 5177-5181.

Blattner FR, Plunkett G, III, Bloch CA*, et al.* (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**: 1453-1462.

Edgar R, Domrachev M & Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucl. Acids Res.* **30**: 207-210.

Gallup JM & Ackermann MR (2006) Addressing fluorogenic real-time qPCR inhibition using the novel custom Excel file system 'Focusfield2-6GallupqPCRSet-upTool-001' to attain consistently high fidelity qPCR reactions. *Biol. Proceed. Online* **8**: 87-152.

Gross CA (1996) Function and regulation of the heat shock proteins. *EcoSal-Escherichia coli and Salmonella: celluar and molecular biology,*(Bock A, Curtiss III R, Kaper JB, Neidhardt FC, Nystrom T, Rudd KE & Squired CL, eds.), pp. ASM Press, Washington, D.C.

Jensen ST & Liu JS (2004) BioOptimizer: a Bayesian scoring function approach to motif discovery. *Bioinformatics* **20**: 1557-1564.

Kanack KJ, Crawford JA, Tatsuno I, Karmali MA & Kaper JB (2005) SepZ/EspZ is secreted and translocated into HeLa cells by the Enteropathogenic *Escherichia coli* type III secretion system. *Infect. Immun.* **73**: 4327-4337.

Liu X, Brutlag DL & Liu JS (2001) BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. *Pac. Symp. Biocomput.* 127-138.

Madsen ML, Nettleton D, Thacker EL, Edwards R & Minion FC (2006) Transcriptional profiling of *Mycoplasma hyopneumoniae* during heat shock using microarrays. *Infect. Immun.* **74**: 160-166.

Morey JS, Ryan JC & Van Dolah FM (2006) Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biol. Proced. Online* **8**: 175-193.

Nataro JP & Kaper JB (1998) Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* **11**: 142-201.

Oneal MJ, Schafer ER, Madsen ML & Minion FC (2008) Global transcriptional analysis of Mycoplasma hyopneumoniae following exposure to norepinephrine. *Microbiology* **154**: 2581-2588.

Perna NT, Plunkett G, Burland V*, et al.* (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**: 529-533.

Raghavachari N, Bao YP, Li G, Xie X & Mèuller UR (2003) Reduction of autofluorescence on DNA microarrays and slide surfaces by treatment with sodium borohydride. *Anal. Biochem.* **312**: 101-105.

Riley LW, Remis RS, Helgerson SD*, et al.* (1983) Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N. Engl. J. Med.* **308**: 681-685.

CHAPTER 3. TRANSCRIPTIONAL RESPONSE OF *ESCHERICHIA COLI* O157:H7 EDL933 DURING ANAEROBIC GROWTH AND EXPOSURE TO RUMEN FLUID

To be submitted to Applied and Environmental Microbiology

Michael D. Carruthers, Nancy Cornick, and F. Chris Minion*

Department of Veterinary Microbiology and Preventive Medicine, and Interdepartmental Microbiology Program, Iowa State University, Ames, IA 50011

*Corresponding Author, fcminion@iastate.edu

Abstract

Escherichia coli O157:H7 encounters two environmental conditions within the ruminant gastrointestinal tract, limiting oxygen and rumen fluid. Both of these conditions may have implications in the general physiology of *E. coli* and its survival during passage through the rumen. The impact of these environmental conditions on the transcriptome of *E. coli* O157:H7 is unclear. To study this further, the transcriptome of *E. coli* O157:H7 strain EDL933 grown under anaerobic conditions was compared to cells exposed to rumen fluid and to cells grown aerobically. Four hundred nineteen genes were differentially expressed during anaerobic growth at $p < 0.014$ with a false discovery rate < 10%. Among the 280 up-regulated genes were the curli pili associated genes *csgBA and csgDEFG*, adherence associated genes *ompA, tdcA,* and *cadA,* heat shock genes *dnaK*, *dnaJ*, *groEL*, and *groES* and the σ^{38} encoding gene *rpoS*. Ten genes, 8 up- and 2 down-regulated, located in *E. coli* O157:H7 virulence-related O-islands were differentially transcribed with a fold change >1.3 suggesting that these genes may play a role in host adaptation. Translational β-galactosidase fusions with *csgBA* confirmed its regulation

during anaerobic conditions. As a result of rumen fluid exposure, 83 genes were differentially transcribed at $p < 0.002$ with a false discovery rate < 10%, 38 up- and 45 down-regulated. There were 11 up-regulated, 11 down-regulated, and 3 opposite-regulated genes common to both conditions. Our analysis indicated that protein synthesis was limited during anaerobiasis and may be further slowed by exposure to rumen fluid.

Introduction

Since its isolation over two decades ago (20), *Escherichia coli* O157:H7 has remained a significant food borne pathogen. The disease caused by *E. coli* O157:H7 in humans is manifested as acute gastroenteritis with symptoms ranging from mild diarrhea to hemorrhagic colitis (23). A potentially fatal sequelae of infection, hemolytic uremic syndrome, is the leading cause of acute renal failure in children (23). Upon entering a host, *E. coli* O157:H7 encounters a variety of conditions, which may influence the organism ability to colonize and/or cause disease.

Once ingested, enteric pathogens are exposed to an oxygen-limiting environment.. *E. coli* has been shown to have differences in growth and metabolic rates as well as an altered transcriptome during anaerobiasis (28, 29). Microorganisms transitioning through these environments must alter transcriptional profiles in order to use simple sugars in anaerobic respiration and fermentation, and they must use alternate terminal electron acceptors. Although the transcriptional response of *E. coli* K-12 (28, 29) and a small number *E. coli* O157:H7 genes (10) to anaerobiasis have been described, the total transcriptome response of *E. coli* O157:H7 under anaerobic conditions has yet to be reported.

Ruminants are considered to be the major environmental reservoir of *E. coli* O157:H7 with organisms persisting in experimentally infected animals up to 2 months or longer (3, 7-9, 13, 16, 31, 41-43). The lower gastrointestinal tract in ruminants, specifically the terminal rectal mucosa, has been implicated as the primary site of *E. coli* O157:H7 colonization (24). To reach this mucosal surface, however, *E. coli* O157:H7 encounters diverse environments during its gastrointestinal tract passage that may impact its fitness for tissue adherence and survival within the lower rectal region.

Research involving rumen carriage of *E. coli* O157:H7 has produced conflicting results. Early studies show that *E. coli* O157:H7 growth is unrestricted in rumen fluid of fasted cattle while growth is inhibited in the rumen fluid of well-fed cattle (26). This sparked research on the effect of ruminant diet, including many different abiotic and biotic factors, in attempts to reduce or eliminate *E. coli* O157:H7 from ruminants. The effect of diet on *E. coli* O157:H7 persistence and shedding continues to be contentious among researchers. The idea that volatile fatty acids produced during carbohydrate fermentation in the rumen inhibits *E. coli* O157:H7 growth was demonstrated *in vitro* (11), but *in vivo* studies have produced conflicting results (18, 33). The effect that pH has on *E. coli* O157:H7 survival in an artificial rumen model appears minimal (32). The effect of supplements such as monesin (39), esculin (12), and others have had limited success in preventing *E. coli* O157:H7 carriage in the rumen and remain to be proven in the field. Results from studies focused on biotic factors have suggested that indigenous rumen microflora inhibit growth or decreased survival of *E. coli* O157:H7 by competitive exclusion (5, 32). Consequently, there is a significant gap in the understanding of the biology of *E. coli* O157:H7 in the rumen environment.

The goal of this study was to measure the transcriptional changes that occur in *E. coli* O157:H7 as it encounters the two major environmental conditions in the rumen, anaerobiasis and rumen fluid exposure, compared to standard aerobic laboratory growth conditions using two-color microarrays. Our results demonstrate that in addition to genes generally involved in anaerobic metabolism and slowed growth, virulence-related genes were also up-regulated. In addition, it was shown that genes not previously identified as being differentially expressed during anaerobiasis in *E. coli* and genes not found within *E. coli* K-12 were differentially expressed in *E. coli* O157:H7 in response to limiting oxygen conditions and rumen fluid exposure.

Materials and Methods

Strains and culture conditions

Escherichia coli strains DH5α (F^- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-) and O157:H7 EDL933

(ATCC 43895) were grown in Luria-Bertani (LB) broth with 0.001% resazurin as an oxygen indicator when necessary. Anaerobic media was purged of oxygen by gentle boiling for 15 min over a flame in a round bottom flask and quickly stoppered. After autoclaving, media was allowed to cool for several hours before being moved into a Bactron anaerobic chamber (Shel Labs, Cornelius, OR). Medium was allowed to equilibrate for 2 days under an atmosphere of 5% CO_2, 5% H_2 and 90% N_2. All manipulations requiring opening of anaerobic tubes or flasks occurred within the anaerobic chamber. Incubation of cultures occurred outside the chamber. Cultures for growth curves were prepared by diluting an overnight culture of *E. coli* O157:H7 1:100 into 5 ml of anaerobic and aerobic LB broth. After incubating overnight cultures at 37°C, they were again diluted 1:100 into 500 ml side-arm flasks with 100 ml LB broth and incubated at 37°C with shaking at 200 rpm. At 15 min intervals, optical densities at 600 nm were taken using a Spectronic 20 (Bausch and Lomb, Rochester, NY) spectrophotometer.

Anaerobic and aerobic cultures were prepared for microarray study by inoculating 5 ml of LB broth with 50 µl of their respective overnight cultures previously diluted to an OD_{600} of 0.6 and incubated for 3 h at 37°C with shaking at 200 rpm. For rumen fluid experiments, overnight anaerobic cultures were centrifuged within the anaerobic chamber for 5 min at 5,000 x *g* and resuspended in 37°C equilibrated, heat clarified anaerobic rumen fluid or anaerobic LB broth, respectively. These cultures were incubated at 37°C for 15 min. Following incubation, 1 ml of each culture was added to 2 ml of RNAprotect Bacteria Reagent (Qiagen, Valencia, CA), pellets were disrupted by vortexing, and samples were incubated at room temperature for 10 min prior to RNA isolation.

Rumen fluid

Rumen fluid was collected from a rumen fluid donor Angus cross breed cow held at the Iowa State University Teaching Hospital and fed a 100% hay diet. Upon collection, rumen fluid was heat clarified by steaming at 121°C for 15 min and allowed to cool to room temperature. Rumen fluid was then centrifuged three times at 10,000 x *g* for 15 min, each time carefully aspirating rumen fluid from the pellet to a new centrifuge bottle.

Rumen fluid was autoclaved for 15 min at 15 psi, 121°C. Sterile heat clarified rumen fluid was stored at 4°C in a stoppered flask until use.

Recombinant DNA techniques

A β-galactosidase fusion with the *csg* promoter was constructed by generating a PCR fragment containing the intergenic region between *csgDEFG* and *csgBA* and inserting it into pLacZY2 (19) to form pLacZY2-P_{csgBA}. The PCR primers used were forward primer csgF (5'-ATTTCTAGAGCTGTCGCCTGCAAA GAAGGTTTA), which contained an XbaI site (underlined) at the 5' end and reverse primer csgR (5'-ATTAAGCTTTTTCATGTTGTCACCCTGGACCTG), which contained the first three codons of *csgB* with a HindIII site (underlined) at the 5' end. The PCR fragment was treated with restriction enzymes, ligated, and cloned directly into XbaI/HindIII digested, dephosphorylated pLacZY2. Plasmids were isolated from DH5α and transformed into *E. coli* O157:H7 EDL933 by electroporation. Transformants were selected on LB agar with 100 μg/ml ampicillin and 20 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside.

β-galactosidase assay

Escherichia coli O157:H7 with and without plasmids pLacZY2 and pLacZY2-P_{csgBA} were grown overnight in LB broth with ampicillin as required at 37°C with shaking. All cultures were diluted 1:100 into LB broth under aerobic and anaerobic conditions, and incubated at 37°C with shaking overnight. These cultures were then diluted 1:100 and grown under their respective oxygen conditions at 37°C with shaking for 3 h for β-galactosidase assays as previously described (21).

RNA isolation

RNA was isolated from samples using the RNeasy Mini Kit (Qiagen). Post-RNA extraction, each sample was treated with DNase (Ambion, Austin, TX) at 37°C for 30 min. Samples were purified and concentrated using Microcon YM-30 columns (Millipore, Billerica, MA), and the quantity and purity were determined using a ND-1000 spectrophotometer (Nanodrop, Wilmington, DE). Samples were determined to be free of contaminating genomic DNA by the absence of a band on a DNA electrophoresis gel

after 30 rounds of PCR. RNA integrity was determined using the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) with all samples measured having a RNA integrity number of 9.0 or higher.

Labeled cDNA generation and hybridization

Target generation, labeling, reaction clean-up, hybridization, pre- and post-hybridization washes were all conducted as previously described (4).

Microarray experimental design

Twelve RNA samples, half from cells grown anaerobically and half from cells grown aerobically, were used for microarray studies in the anaerobic to aerobic comparison. Fourteen RNA samples, half from anaerobically grown cells treated with rumen fluid and half treated with fresh anaerobic LB broth, were used for the rumen fluid treatment microarray studies. Two RNA samples, one from each treatment, were paired randomly for hybridization on a two-color microarray slide for a total of 6 hybridizations for the anaerobic to aerobic comparison and 7 hybridizations for the rumen fluid treatment comparison. Dye assignments to control and treated samples were reversed for three of the arrays to account for differences in fluorescence intensity of the dyes (dye bias).

Image acquisition and data analysis

Scanning, image segmentation, and normalization were conducted as previously described (4). The normalized values for duplicate spots were averaged within each array to produce one normalized measure of fluorescence for each probe sequence and each of the experimental units. Mixed model analysis was conducted as previously described (4). Cluster of Orthologous Groups of proteins (COGs) information was obtained from NCBI (http://www.ncbi.nlm.nih.gov).

Validation of microarray data.

To confirm significant transcriptional differences due to treatment, semi-quantitative real time polymerase chain reaction (qRT-PCR) was performed on five genes in the anaerobic to aerobic comparison (*cheY*, *csgA*, *dnaK*, *glpD*, *groEL*) and six genes in

the rumen fluid treatment study (*csgA*, *dnaK*, *malF*, *ompA*, *rpoS*, *rpsG*) shown to have significant transcriptional differences as a result of treatment as determined by microarrays. Six RNA samples from each treatment were randomly chosen for use in qRT-PCR and were represented on each plate in duplicate. Four point standard curves in duplicate were used for each primer set including the reference gene. The Express One-Step SYBR GreenER kit (Invitrogen) was used for qRT-PCR with the Mx3005P QPCR System (Stratagene, La Jolla, CA). The method described in Gallup et al. was used in qRT-PCR experimental design, and the ISU equation was used to calculate fold change between treatment and control samples (14). The t test was used to determine if the difference between treatment and control were significant. qRT-PCR primers used in this study were designed using PrimerQuest (Integrated DNA Technologies, Coralville, IA), synthesized by IDT and are described in Table 3.1.

Microarray data accession

The dataset for the anaerobic to aerobic comparison can be accessed from the National Center for Biological Informatics Gene Expression Omnibus using Series accession number GSE13301 (http://www.ncbi.nlm.nih.gov/geo/). The dataset for the rumen fluid experiment has Series accession number GSE16747 (http://www.ncbi.nlm.nih.gov/geo/).

Table 3.1. qRT-PCR primers used to validate microarray results

Gene	Sequence	Tm °C
cheY-F	CAA CAT GGA TGG CCT GGA ATT GCT	60.2
cheY-R	CGC CGC AGC AAT GAT GTT CTC TTT	60.4
csgA-F	TGG TGG TAA CTC TGC ACT TGC TCT	60.3
csgA-R	TGT CAT CTG AGC CCT GAC CAA CAT	60.2
dnaK-F	AGC TGG CAG GAC AAA GTC GAA GTA	60.2
dnaK-R	CAT TCG CGC CAT CAC CTC TTC AAT	59.9
fhlA-F	TTC AAA CCT TGT GCC TGT TAC CGC	60.2
fhlA-R	GTT CGG CAA TCT GGC GCA GTA AAT	60.3
glpD-F	TGA AAC AGT TCT TCG ACG ACG GGA	60.2
glpD-R	GCA CCA CAA TAT GGC TGC CTT TGA	60.2
groEL-F	GTG GGT ATC AAA GTT GCA CTG CGT	60.0
groEL-R	TTT GGT TGG GTC CAG GAT ACC CAT	60.2
malF-F	ATT CAA CTG TTA ACC AAC GGC GGC	60.4
malF-R	TGG CTT TCA GGT TCA CTA TCG CCA	60.5
malF-F	ATT CAA CTG TTA ACC AAC GGC GGC	60.4
rpoS-F	TCT CAA CAT ACG CAA CCT GGT GGA	60.3
rpoS-R	TGG TTC ATG GTC CAG CTT ATG GGA	60.2
rpsG-F	TGG CAA TGC GTT GGA TCG TTG AAG	60.4
rpsG-R	TGC AGC ATC AGA AAG TTC GTT CGC	60.2
ompA-F	TAC CAG TGG ACC AAC AAC ATC GGT	60.3
ompA-R	AAG TGC TTG GTC TGT ACT TCC GGT	60.3

Results

Anaerobic vs. aerobic microarrays

To determine the effect of anaerobiasis on the transcriptome of *E. coli* O157:H7 EDL933, microarrays were used to compare transcript levels within planktonic cultures subjected to anaerobic conditions with those of cultures grown aerobically. Cultures were approximately at the same density and phase of growth (Fig. 3.1). Extraction of total RNA from approximately 6.0 x 10^8 cells yielded a minimum of 25 µg of total RNA post-DNase treatment. Equal amounts of paired aerobic and anaerobic total RNA from all samples was subjected to cDNA synthesis and dye coupling to allow for an equal number of dye swaps. Data from each of 6 replicates were used in the statistical analysis.

Statistical analysis indicated that 419 genes demonstrated transcriptional differences with a p-value < 0.014 and an estimated false discovery rate (FDR) < 10%.

Significance and differences in transcript levels for all genes are depicted as a volcano plot (Fig. 3.2). Of the 419 genes differentially expressed, 280 genes were up regulated while 139 genes were down regulated. Selected differentially expressed genes are listed in Tables 3.2 and 3.3, respectively. The remaining genes are listed in supplemental table 3.S1 (Appendix B). These genes cover 21 clusters of orthologous groups of proteins (COGs) as shown in Fig. 3.3. There were 82 hypothetical genes, 62 up- and 20 down-regulated.

Five genes, 3 up-regulated and 2 down-regulated, were chosen for validation using qRT-PCR. *phoP* was used for the control gene since it did not show differential expression in the microarray study. In every case, the qRT-PCR results corroborated the microarray results with respect to direction of differential expression (Fig. 3.4). The degree of transcript difference as measured by qRT-PCR was greater than measured by microarray as has been shown previously (22).

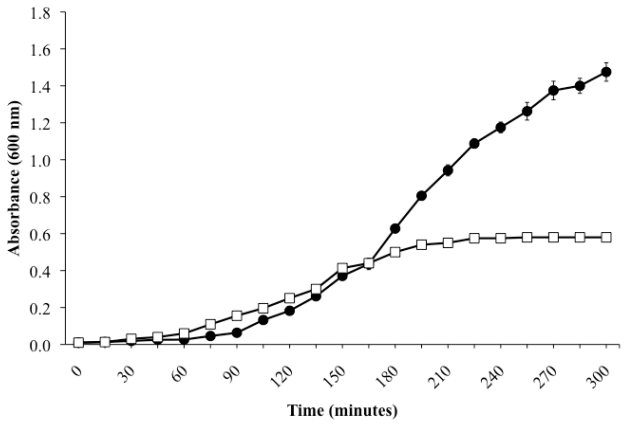

Figure 3.1. Temporal growth of *E. coli* O157:H7 EDL933 in LB broth under anaerobic and normal atmospheric oxygen conditions as measured by OD600. Data represent the mean and standard error of three replicates. Circle indicates aerobic growth and open square indicates anaerobic growth.

Table 3.2. Select up-regulated transcripts as a result of anaerobiasis at FDR < 10%

Gene	Locus	FC[1]	FC[2]	*P*-value	Product/Function
ackA	Z3558	1.9		0.00001	Acetate propionate kinase
adhE	Z2016	1.6		0.00043	CoA-linked acetaldehyde dehydrogenase
cadA	Z5734	1.3		0.00314	Lysine decarboxylase 1
csgA	Z1676	1.9	11.7	0.00054	Curlin major subunit
csgB	Z1675	1.4		0.00018	Minor curlin subunit precursor
csgC	Z1677	1.5		0.00012	Putative curli production protein
csgD	Z1673	1.2		0.01004	Putative 2-component transcriptional regulator
csgE	Z1672	1.3		0.00251	Curli assembly transport component
csgF	Z1671	1.5		0.01254	Curli assembly transport component
csgG	Z1670	1.4		0.00400	Curli assembly transport component
dnaJ	Z0015	1.3		0.00083	Chaperone
dnaK	Z0014	2.0	6.7	0.00031	Molecular chaperone
fdhF	Z5678	1.4		0.00002	Formate dehydrogenase
groEL	Z5748	1.5	3.3	0.00011	Chaperonin
groES	Z5747	1.7		0.00011	Co-chaperonin
hycE	Z4029	1.2		0.00099	Large subunit of hydrogenase 3 p
hycG	Z4027	1.8		0.00054	Hydrogenase
ompA	Z1307	1.4		0.00087	Outer membrane protein A
pta	Z3559	2.1		0.00001	Phosphate acetyltransferase
rpoS	Z4049	2.0		0.00002	RNA polymerase sigma factor
tdcA	Z4470	1.2		0.00570	Transcriptional activator of *tdc* operon
tdcE	Z4466	2.7		0.00032	Probable formate acetyltransferase 3
Z0615*	Z0615	1.2		0.00055	Putative RTX family exoprotein
Z0634*	Z0634	1.5		0.00002	Putative cytoplasmic membrane export protein
Z0968*	Z0968	2.2		0.00014	Hypothetical protein
Z1444*	Z1444	1.5		0.00594	Putative serine threonine kinase within bacteriophage BP-933W
Z1453*	Z1453	1.4		0.00886	Hypothetical protein
Z1498*	Z1498	1.6		0.00474	Hypothetical protein
Z3332*	Z3332	1.5		0.00085	Hypothetical protein
Z6010*	Z6010	1.4		0.01025	Hypothetical protein

[1] FC = fold change determined by microarray
[2] FC = fold change determined by qRT-PCR
* Encoded within an O-island

Table 3.3. Select down-regulated transcripts as a result of anaerobiasis at FDR < 10%

Gene	Locus	FC[1]	FC[2]	*P*-value	Product/Function
cheY	Z2936	-1.1	-2.2	0.0107126	Chemotaxis regulator
fliC	Z3013	-1.6		0.0002183	Flagellar biosynthesis protein
glpD	Z4786	-2.5	-15.2	0.0000002	sn-Glycerol-3-phosphate dehydrogenase
rplB	Z4688	-1.4		0.0001761	50S ribosomal protein L2
rplC	Z4691	-1.8		0.0000113	50S ribosomal protein L3
rplD	Z4690	-2.1		0.0000161	50S ribosomal protein L4
rplE	Z4678	-1.4		0.0000464	50S ribosomal protein L5
rplO	Z4671	-1.2		0.0000110	50S ribosomal protein L15
rplP	Z4684	-1.8		0.0000121	50S ribosomal protein L16
rplR	Z4674	-1.3		0.0013214	50S ribosomal protein L18
rplV	Z4686	-2.0		0.0000134	50S ribosomal protein L22
rplX	Z4679	-1.5		0.0001732	50S ribosomal protein L24
rpmC	Z4683	-1.9		0.0000010	50S ribosomal protein L29
rpsD	Z4666	-1.3		0.0085753	30S ribosomal protein S4
rpsE	Z4673	-2.2		0.0000338	30S ribosomal protein S5
rpsG	Z4699	-2.3		0.0000146	30S ribosomal protein S7
rpsH	Z4676	-1.6		0.0000291	30S ribosomal protein S8
rpsI	Z4588	-1.6		0.0004589	30S ribosomal protein S9
rpsJ	Z4692	-1.6		0.0000190	30S ribosomal protein S10
rpsK	Z4667	-1.8		0.0000190	30S ribosomal protein S11
rpsM	Z4668	-1.2		0.0025847	30S ribosomal protein S13
rpsQ	Z4681	-1.2		0.0007669	30S ribosomal protein S17
rpsS	Z4687	-1.9		0.0000177	30S ribosomal protein S19
*wbdP**	Z3199	-1.3		0.0052208	Glycosyl transferase
yaeE	Z0210	-1.4		0.0013091	Putative transport permease protein
Z0655*	Z0655	-1.4		0.0000215	Hypothetical protein

[1]FC = microarray fold change
[2]FC = qRT-PCR fold change
* Encoded within an O-island

Rumen fluid microarrays

To compare the effect of rumen fluid treatment on anaerobically grown *E. coli* O157:H7 EDL933, microarrays were used to compare transcript levels of planktonic anaerobic cultures resuspended in anaerobic rumen fluid to identical cultures resuspended in fresh anaerobic LB broth. Data from each of 7 replicates were used in the statistical analysis. The mixed model analysis indicated that 83 genes demonstrated transcriptional differences with a p-value < 0.002 and an estimated false discovery rate (FDR) < 10%.

Significance and differences in transcript levels for all genes are depicted as a volcano plot (Fig. 3.2). Of the 83 genes differentially expressed, 38 genes were up regulated while 45 genes were down regulated (Table 3.S2)(Appendix C). These genes cover 21 clusters of orthologous groups of proteins (COGs) as shown in Figure 3.3. Differentially transcribed genes common to both microarray studies are listed in Table 3.4.

Six up-regulated and two down-regulated genes were chosen for validation using qRT-PCR. *phoP* was used as the reference gene as it did not show differential expression in the microarray study. In every case the qRT-PCR results corroborated the microarray results with respect to direction of differential expression (Fig. 3.4).

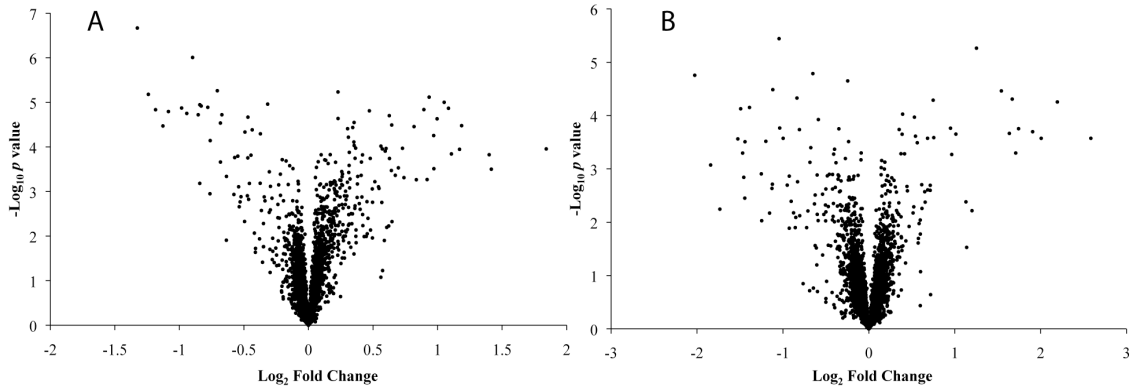

Figure 3.2. Volcano plots of transcriptional differences of *E. coli* O157:H7 during anaerobic growth (A) and rumen fluid treatment (B). Differences are plotted as Log_2 Fold Change vs.-Log_{10} p-value.

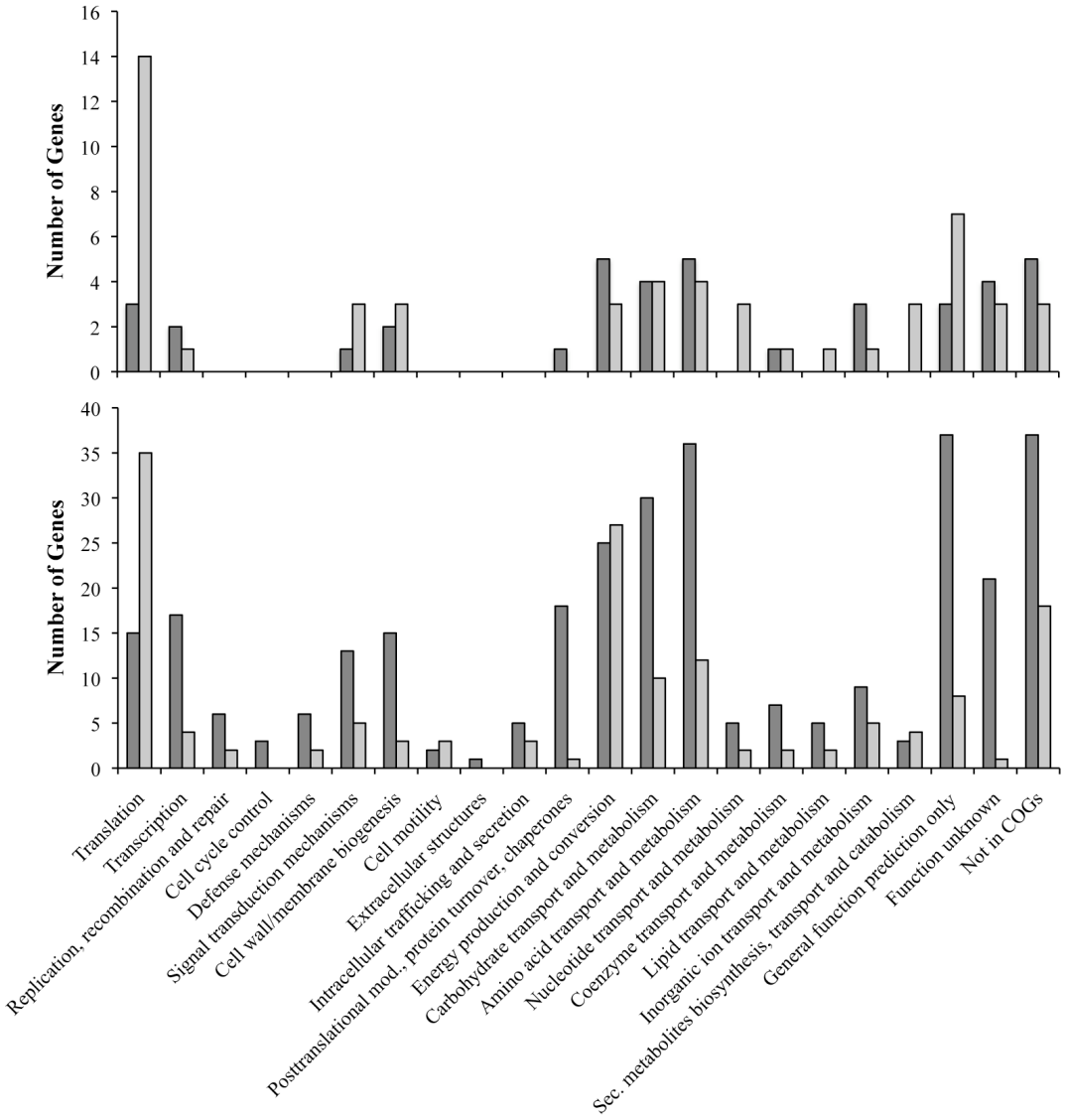

Figure 3.3. Differentially expressed transcripts of Clusters of Orthologous Groups of proteins. Dark gray bars indicate up-regulated genes; light gray bars indicate down-regulated genes. Upper panel is rumen fluid exposure; lower panel is anaerobic growth comparison. Clusters are indicated along the horizontal axis; numbers of genes along the vertical axis.

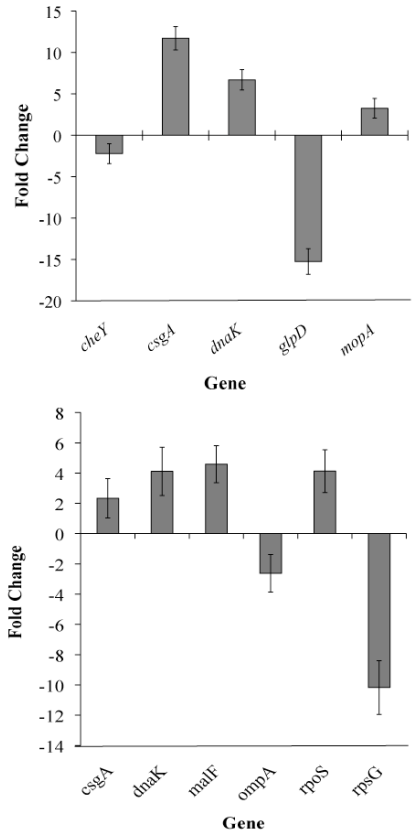

Figure 3.4. Validation of microarray results by qRT-PCR. All genes were differentially expressed at $p < 0.01$. Upper panel represents RNA from anaerobic experiment; lower panel represents data from rumen fluid exposure experiment. Data represents mean of fold change ± standard error.

Table 3.4. Significantly differentially expressed transcripts common to both microarray experiments at FDR < 10%

Gene	Locus	FC[1]	FC[2]	FC[3]	FC[4]	Product/Function
Up-regulated						
csgA	Z1676	1.9	11.3	1.5	2.3	Curlin major subunit
cysK	Z3680	1.3		4.6		Cysteine synthase A
dnaK	Z0014	2.0	6.7	2.0	4.1	Molecular chaperone
gadA	Z4930	3.6		3.2		Glutamate decarboxylase isozyme
hfq	Z5779	1.7		1.3		RNA-binding protein
lysU	Z5732	1.9		1.5		Lysyl-tRNA synthetase
nlpD	Z4050	2.1		3.1		Lipoprotein
rpoS	Z4049	2.0		2.0	4.1	RNA polymerase sigma factor
uspE	Z2435	1.2		3.3		Universal stress protein
yeaU	Z2843	1.2		4.0		Putative tartrate dehydrogenase
yhaR	Z4465	2.6		3.4		Hypothetical protein
Down-regulated						
fusA	Z4698	-2.4		-2.7		Elongation factor G
rplC	Z4691	-1.8		-2.8		50S ribosomal protein L3
rplD	Z4690	-2.1		-2.7		50S ribosomal protein L4
rplP	Z4684	-1.8		-1.9		50S ribosomal protein L16
rplV	Z4686	-2.0		-2.6		50S ribosomal protein L22
rpsC	Z4685	-1.3		-1.8		30S ribosomal protein S3
rpsG	Z4699	-2.3		-3.6	-10.2	30S ribosomal protein S7
rpsH	Z4676	-1.6		-1.9		30S ribosomal protein S8
rpsK	Z4667	-1.8		-2.8		30S ribosomal protein S11
rpsS	Z4687	-1.9		-2.4		30S ribosomal protein S19
ybbA	Z0648	-1.6		-2.2		Putative ABC transporter ATP-binding protein
Opposite regulation						
pta	Z3559	2.1		-4.1		Phosphate acetyltransferase
ompA	Z1307	1.4		-1.6	-2.6	Outer membrane protein A
hisS	Z3777	1.2		-1.4		Histidyl-tRNA synthetase

[1]FC = microarray fold change of anaerobic to aerobic comparison
[2]FC = qRT-PCR fold change of anaerobic to aerobic comparison
[3]FC = microarray fold change of rumen fluid to LB broth comparison
[4]FC = qRT-PCR fold change of rumen fluid to LB broth comparison

β-galactosidase assay

To confirm findings of curli up-regulation during anaerobic growth, a translational fusion to β-galactosidase was constructed with DNA sequences from the

promoter of *csgBA* and the first few amino acids of *csgB*. Strains were grown under anaerobic and aerobic conditions, and β-galactosidase assays were performed. Translation from the promoter of *csgBA* was 1.8 fold higher under anaerobic conditions compared to aerobic conditions at a *p* value of 8.5×10^{-6} as detected by β-galactosidase activity (Fig. 3.5).

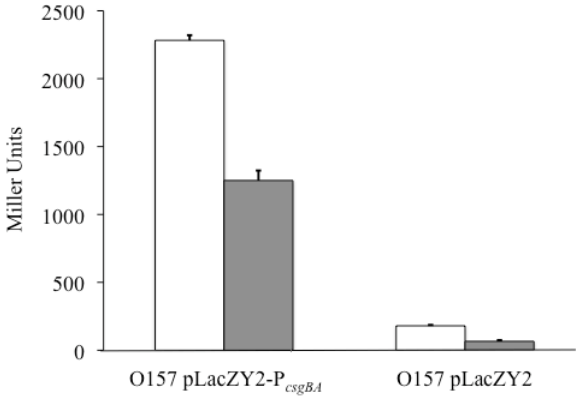

Figure 3.5. β-galactosidase assay results of a *csgBA* fusion under anaerobic and aerobic conditions. Anaerobic condition represented as white bars and aerobic condition as gray bars. pLacZY2 is the negative control.

Discussion

The bovine gastrointestinal tract is the major reservoir for *E. coli* O157:H7. There the organism colonizes the lower mucosal epithelium and is shed intermittently into the environment (7, 16). Contaminated feces is the main source for the spread of *E. coli* O157:H7 into the human food supply either directly through feces contamination of meat surfaces during processing (40) or through feces contamination of water sources used for irrigation or the washing of produce following harvest (30). Naïve cattle are frequently exposed to *E. coli* O157:H7 in the cattle feedlot environment primarily from penmates that are transiently shedding the organism. It is also thought this is the mechanism by which *E. coli* maintains itself in cattle populations through a continual re-infection model.

As it traverses the gastrointestinal tract, *E. coli* O157:H7 encounters a wide variety of environments. Its responses to these environmental changes are not well characterized, and it is not known if these changes increase its fitness for the intestinal tract environment, thereby enhancing its survival along the colonic mucosal surface.

To better understand the extent of transcriptional changes that occur in *E. coli* O157:H7 during its passage through the bovine gastrointestinal tract, microarray studies were performed. These studies were designed to mimic two environmental changes that *E. coli* O157:H7 encounters early in its path through the bovine rumen, anaerobiasis and exposure to rumen fluid under anaerobic conditions. The study design was such that we separated the two environmental changes, one involving an aerobic to anaerobic shift and one involving an anaerobic laboratory media (LB broth) to anaerobic rumen fluid shift, all at 37°C and with minimal pH changes. We separated the two environmental shifts to better understand the flow of transcriptional changes in *E. coli* O157:H7 as it enters the bovine rumen.

Our data clearly shows that *E. coli* O157:H7 significantly alters its steady state transcription patterns in response to anaerobiasis. This result was not surprising since *E. coli* K-12 also undergoes transcriptional changes during the transition from oxygen replete to oxygen limiting conditions (28). Genes annotated as having hypothetical protein products comprise a significant portion of the differentially transcribed genes under anaerobiasis (19.5%) and following rumen fluid exposure (16.9%)(Table 3.S1)(Appendix B). Seventeen of the genes found to be down-regulated during anaerobiasis were found to be differentially expressed by Salmon *et al*. in *E. coli* K-12 as well, but only six of the genes found by our study to be up-regulated during anaerobiasis were identified by Salmon *et al*. (28). These differences may be attributed to differences in experimental conditions, array platforms, and data analysis or to differences in the genetic backgrounds. Further studies will be needed to differentiate these possibilities.

A large number of metabolic genes were differentially transcribed between aerobic and anaerobic conditions. This result was expected, as there are a number of differences in the metabolism of *E. coli* in the presence or absence of oxygen and thus

differential expression would be needed to accommodate those changes. Many ribosomal associated genes including the 30S ribosomal associated transcript *rpsMKD-rpoA-rplQ* and the 50S ribosomal associated transcript *rpsJ-rplCDWB-rpsS-rplB-rpsC-rplP-rpmC-rpsQ* were down-regulated, which was previously seen in *E. coli* as a result of anaerobiasis (28). This was anticipated; anaerobic growth rates are slower than aerobic growth rates, and thus the need for ribosomes and translational associated proteins is reduced. This was also seen during treatment with anaerobic rumen fluid indicating that growth rate and protein synthesis is further slowed upon entry into the rumen. This is thought to be an energy conservation response to the environmental changes as is the down-regulation of aerobic respiration. The anaerobic response has 21 down-regulated genes annotated as having aerobic respiration functions. Genes encoding the cytochrome O ubiquinol oxidase and a NADH dehydrogenase, *cyoABCDE* and *nuoABCDEFGHIJKLMN* respectively were down-regulated. There were 9 genes up-regulated under anaerobic conditions that are involved in mixed acid fermentation including *pta, ackA, adhE, fdhF, tdcE, hycE, hycF, hycB,* and *hycG.* Sixteen genes up-regulated by RpoN and 27 genes up-regulated by RpoS (http://www.ecocyc.org) were also up-regulated in our study as a result of anaerobiasis.

Escherichia coli O157:H7 EDL933 produces curli fimbrae under certain environmental conditions (37). Expression of curli genes has been shown to be under the control of CsgD and RpoS (1, 17, 25), both of which were up-regulated in our study. Increased expression of *csgD* leads to increased virulence in mice and increased invasion of Hep-2 cells by *E. coli* O157:H7 (38). All of the genes found within the divergent *E. coli* O157:H7 EDL933 *csg* loci, *csgBA* and *csgDEFG*, were up-regulated under anaerobic conditions which has been previously described in *E. coli* O157:H7 (10) but not validated translationally. Exposure to rumen fluid resulted in further up-regulation of *csgA*. These results suggest that oxygen-poor environments such as the gastrointestinal tract of mammalian hosts induce genes involved in producing attachment structures. This type of oxygen sensitive regulation has been previously described in *Salmonella* (15). Regulating production of curli only when necessary, i.e., for attachment inside the host, would be

energetically favorable and would limit antigenic exposure. Our *csgB* translational fusion with β-galactosidase demonstrates that these proteins are produced at higher levels after 3 h of growth under anaerobiasis. The possible impact of these differences on colonization within the ruminant and pathogenesis within humans is not known, but recent studies show that CsgA enhances biofilm formation and HEp-2 cell invasion by *E. coli* O157:H7 (27, 36).

Three additional genes that have been implicated in adherence, *cadA, tdcA* and *ompA* were also up-regulated as a result of anaerobiasis (34, 35). Mutations of *E. coli* O157:H7 *tdcA* demonstrate an increase in adherence to HeLa and Caco-2 cells compared to wild-type by negative regulation of *ompA* (35). In contrast, the *E. coli* O157:H7 genotypes Δ*ompA* and Δ*ompA:*Δ*tdcA* both show a striking reduction in adherence to alfalfa and seed coats (34). Our studies demonstrate that both *tdcA* and *ompA* were up-regulated as a result of anaerobiasis. The significance of both *tdcA* and *ompA* up-regulation is not known, but up-regulation of *tdcA* is not enough to prevent up-regulation of *ompA* under anaerobic conditions. Exposure of *E. coli* O157:H7 to anaerobic rumen fluid, however, resulted in down-regulation of *ompA,* which may negate the effects of anaerobiasis.

Ten differentially expressed genes are located in *E. coli* O157:H7 O-islands, 8 up- and 2 down-regulated (Tables 2 and 3). Five of these up-regulated genes and one down-regulated gene were hypothetical. The three remaining up-regulated O-island genes were a cytoplasmic export protein (Z0634), a serine/threonine kinase encoded on bacteriophage BP-933W (Z1444), and a putative RTX family exoprotein (Z0615).

Gene *wbdP*, which encodes a glycosyl transferase, was down-regulated in our studies during anaerobiasis. It is on the same transcriptional unit as *manC* and *wxy*, which encode mannose-1-P guanosyltransferase and O antigen polymerase, respectively, and are involved in the synthesis of surface polysaccharides. It has been shown that *E. coli* O157:H7 mutants deficient in the expression LPS are more adherent to Hep-2 and HeLa cells than wild-type (2, 6). This indicates that *E. coli* O157:H7 may be down-regulating

genes involved in lengthening LPS, which may in turn allow for uncovering of afimbrial adhesins.

This study demonstrates that *E. coli* O157:H7 up-regulates fermentative metabolism and anaerobic respiration when it encounters rumen mimicking environmental conditions. Also up-regulated were select virulence factors associated primarily with cell adhesion. The impact of rumen fluid on transcription was minimal, but in some instances the transcription changes could increase fitness for colonization and survival. This pattern of regulation may be beneficial by using limiting oxygen as a queue for up-regulating energy intensive, virulence-related processes such as adherence factors in low energy environments, i.e., anaerobiasis. Many of the studies on virulence factors and their regulation are conducted in oxygen replete environments such as the assay of cell adherence with *in vitro* cell cultures. Our studies show there is a need for the study of virulence factors under environments that are more similar to anaerobic *in vivo* conditions.

Acknowledgements

We would like to thank Dr. Thaddeus Stanton and Sam Humphrey of the National Animal Disease Center for their advice regarding anaerobic work. We would also like to thank Dr. Greg Phillips for providing pLacZY2. This study was funded in part by a contracts to FCM and NC from the United States Department of Agriculture (Specific Cooperative Agreements 58-3625-2-127 and 58-3625-4-173).

References

1. **Arnqvist, A., A. Olsen, and S. Normark.** 1994. Sigma S-dependent growth-phase induction of the *csgBA* promoter in *Escherichia coli* can be achieved *in vivo* by sigma 70 in the absence of the nucleoid-associated protein H-NS. Mol. Microbiol. **13:**1021-1032.

2. **Bilge, S. S., J. C. Vary, Jr., S. F. Dowell, and P. I. Tarr.** 1996. Role of the *Escherichia coli* O157:H7 O side chain in adherence and analysis of an *rfb* locus. Infect. Immun. **64:**4795-4801.

3. **Brown, C. A., B. G. Harmon, T. Zhao, and M. P. Doyle.** 1997. Experimental *Escherichia coli* O157:H7 carriage in calves. Appl. Environ. Microbiol. **63:**27-32.

4. **Carruthers, M. D., and C. Minion.** 2009. Transcriptome analysis of *Escherichia coli* O157:H7 EDL933 during heat shock. FEMS Microbiol. Lett. **295:**96-102.

5. **Chaucheyras-Durand, F., J. Madic, F. Doudin, and C. Martin.** 2006. Biotic and Abiotic Factors Influencing *In Vitro* Growth of *Escherichia coli* O157:H7 in Ruminant Digestive Contents. Appl. Environ. Microbiol. **72:**4136-4142.

6. **Cockerill 3rd, F., G. Beebakhee, R. Soni, and P. Sherman.** 1996. Polysaccharide side chains are not required for attaching and effacing adhesion of *Escherichia coli* O157:H7. Infect. Immun. **64:**3196-3200.

7. **Cray, W. C., Jr., and H. W. Moon.** 1995. Experimental infection of calves and adult cattle with *Escherichia coli* O157:H7. Appl. Environ. Microbiol. **61:**1586-1590.

8. **Dean-Nystrom, E. A., B. T. Bosworth, W. C. Cray, Jr., and H. W. Moon.** 1997. Pathogenicity of *Escherichia coli* O157:H7 in the intestines of neonatal calves. Infect. Immun. **65:**1842-1848.

9. **Dean-Nystrom, E. A., B. T. Bosworth, and H. W. Moon.** 1997. Pathogenesis of O157:H7 *Escherichia coli* infection in neonatal calves. Adv. Exp. Med. Biol. **412:**47-51.

10. **Dowd, S. E., and H. Ishizaki.** 2006. Microarray based comparison of two *Escherichia coli* O157:H7 lineages. BMC Microbiol. **6:**30.

11. **Duncan, S. H., H. J. Flint, and C. S. Stewart.** 1998. Inhibitory activity of gut bacteria against *Escherichia coli* O157 mediated by dietary plant metabolites. FEMS Microbiol. Lett. **164:**283-238.

12. **Duncan, S. H., E. C. Leitch, K. N. Stanley, A. J. Richardson, R. A. Laven, H. J. Flint, and C. S. Stewart.** 2004. Effects of esculin and esculetin on the survival of *Escherichia coli* O157 in human faecal slurries, continuous-flow simulations of the rumen and colon and in calves. Br. J. Nutr. **91:**749-755.

13. **Dziva, F., P. M. van Diemen, M. P. Stevens, A. J. Smith, and T. S. Wallis.** 2004. Identification of *Escherichia coli* O157:H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis. Microbiol. **150:**3631-3645.

14. **Gallup, J. M., and M. R. Ackermann.** 2006. Addressing fluorogenic real-time qPCR inhibition using the novel custom Excel file system 'Focusfield2-6GallupqPCRSet-upTool-001' to attain consistently high fidelity qPCR reactions. Biol. Proceed. Online. **8:**87-152.

15. **Gerstel, U., and U. Römling.** 2003. The *csgD* promoter, a control unit for biofilm formation in *Salmonella typhimurium*. Res. Microbiol. **154:**659-667.

16. **Grauke, L. J., I. T. Kudva, J. W. Yoon, C. W. Hunt, C. J. Williams, and C. J. Hovde.** 2002. Gastrointestinal tract location of *Escherichia coli* O157:H7 in ruminants. Appl. Environ. Microbiol. **68:**2269-2277.

17. **Hammar, M., A. Arnqvist, Z. Bian, A. Olsen, and S. Normark.** 1995. Expression of two *csg* operons is required for production of fibronectin- and congo red-binding curli polymers in *Escherichia coli* K-12. Mol Microbiol **18:**661-670.

18. **Harmon, B. G., C. A. Brown, S. Tkalcic, P. O. Mueller, A. Parks, A. V. Jain, T. Zhao, and M. P. Doyle.** 1999. Fecal shedding and rumen growth of *Escherichia coli* O157:H7 in fasted calves. J. Food. Prot. **62:**574-579.

19. **Jain, C.** 1993. New improved *lacZ* gene fusion vectors. Gene. **133:**99-102.

20. **Kaper, J. B.** 1998. Enterohemorrhagic *Escherichia coli*. Cur. Opin. Microbiol. **1:**103-108.

21. **Miller, J. H.** 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

22. **Morey, J. S., J. C. Ryan, and F. M. Van Dolah.** 2006. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. Biol. Proced. Online. **8:**175-193.

23. **Nataro, J. P., and J. B. Kaper.** 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11:**142-201.

24. **Naylor, S. W., J. C. Low, T. E. Besser, A. Mahajan, G. J. Gunn, M. C. Pearce, I. J. McKendrick, D. G. Smith, and D. L. Gally.** 2003. Lymphoid follicle-dense mucosa at the terminal rectum is the principal site of colonization of enterohemorrhagic *Escherichia coli* O157:H7 in the bovine host. Infect. Immun. **71:**1505-1512.

25. **Olsen, A., A. Arnqvist, M. Hammar, and S. Normark.** 1993. Environmental regulation of curli production in *Escherichia coli*. Infect. Agents. Dis. **2:**272-274.

26. **Rasmussen, M. A., W. C. Cray, Jr., T. A. Casey, and S. C. Whipp.** 1993. Rumen contents as a reservoir of enterohemorrhagic *Escherichia coli*. FEMS Microbiol. Lett. **114:**79-84.

27. **Saldana, Z., J. Xicohtencatl-Cortes, F. Avelino, A. D. Phillips, J. B. Kaper, J. L. Puente, and J. A. Giron.** 2009. Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing *Escherichia coli* and identification of Fis as a negative regulator of curli. Environ. Microbiol. **11:**992-1006.

28. **Salmon, K., S. P. Hung, K. Mekjian, P. Baldi, G. W. Hatfield, and R. P. Gunsalus.** 2003. Global gene expression profiling in *Escherichia coli* K12. The effects of oxygen availability and FNR. J. Biol. Chem. **278:**29837-29855.

29. **Salmon, K. A., S.-p. Hung, N. R. Steffen, R. Krupp, P. Baldi, G. W. Hatfield, and R. P. Gunsalus.** 2005. Global Gene Expression Profiling in *Escherichia coli* K12: Effects of Oxygen Availability and ArcA. J. Biol. Chem. **280:**15084-15096.

30. **Sivapalasingam, S., C. R. Friedman, L. Cohen, and R. V. Tauxe.** 2004. Fresh produce: a growing cause of outbreaks of foodborne illness in the United States, 1973 through 1997. J. Food. Prot. **67:**2342-53.

31. **Stevens, M. P., P. M. van Diemen, G. Frankel, A. D. Phillips, and T. S. Wallis.** 2002. *Efa1* Influences Colonization of the Bovine Intestine by Shiga Toxin-Producing *Escherichia coli* Serotypes O5 and O111. Infect. Immun. **70:**5158-5166.

32. **Thran, B. H., H. S. Hussein, D. Redelman, and G. C. Fernandez.** 2003. Influence of pH treatments on survival of *Escherichia coli* O157:H7 in continuous cultures of rumen contents. Exp. Biol. Med. (Maywood) **228:**365-369.

33. **Tkalcic, S., C. A. Brown, B. G. Harmon, A. V. Jain, E. P. Mueller, A. Parks, K. L. Jacobsen, S. A. Martin, T. Zhao, and M. P. Doyle.** 2000. Effects of diet on rumen proliferation and fecal shedding of *Escherichia coil* O157:H7 in calves. J. Food. Prot. **63:**1630-1636.

34. **Torres, A. G., C. Jeter, W. Langley, and A. G. Matthysse.** 2005. Differential Binding of *Escherichia coli* O157:H7 to Alfalfa, Human Epithelial Cells, and Plastic Is Mediated by a Variety of Surface Structures. Appl. Environ. Microbiol. **71:**8008-8015.

35. **Torres, A. G., and J. B. Kaper.** 2003. Multiple Elements Controlling Adherence of Enterohemorrhagic *Escherichia coli* O157:H7 to HeLa Cells. Infect. Immun. **71:**4985-4995.

36. **Uhlich, G. A., N. W. Gunther 4th, D. O. Bayles, and D. A. Mosier.** 2009. The CsgA and Lpp proteins of an *Escherichia coli* O157:H7 strain affect HEp-2 cell invasion, motility, and biofilm formation. Infect. Immun. **77:**1543-52.

37. **Uhlich, G. A., J. E. Keen, and R. O. Elder.** 2001. Mutations in the *csgD* Promoter Associated with Variations in Curli Expression in Certain Strains of *Escherichia coli* O157:H7. Appl. Environ. Microbiol. **67:**2367-2370.

38. **Uhlich, G. A., J. E. Keen, and R. O. Elder.** 2002. Variations in the *csgD* Promoter of *Escherichia coli* O157:H7 Associated with Increased Virulence in Mice and Increased Invasion of HEp-2 Cells. Infect. Immun. **70:**395-399.

39. **Van Baale, M. J., J. M. Sargeant, D. P. Gnad, B. M. DeBey, K. F. Lechtenberg, and T. G. Nagaraja.** 2004. Effect of forage or grain diets with or without monensin on ruminal persistence and fecal *Escherichia coli* O157:H7 in cattle. Appl. Environ. Microbiol. **70:**5336-5342.

40. **Vugia, D., A. Cronquist, J. Hadler, M. Tobin-D'Angelo, D. Blythe, K. Smith, S. Lathrop, D. Morse, P. Cieslak, F. Jones, K. G. Holt, J. J. Guzewich, O. L. Henao, E. Scallan, F. J. Angulo, P. M. Griffin, and R. V. Tauxe.** 2007. Preliminary FoodNet data on the incidence of infection with pathogens transmitted commonly through food --- 10 states. Morbid. Mortal. Weekly Rep. **56:**336-339.

41. **Wales, A. D., G. R. Pearson, A. M. Skuse, J. M. Roe, C. M. Hayes, A. L. Cookson, and M. J. Woodward.** 2001. Attaching and effacing lesions caused by *Escherichia coli* O157:H7 in experimentally inoculated neonatal lambs. J. Med. Microbiol. **50:**752-758.

42. **Woodward, M. J., A. Best, K. A. Sprigings, G. R. Pearson, A. M. Skuse, A. Wales, C. M. Hayes, J. M. Roe, J. C. Low, and R. M. La Ragione.** 2003. Non-toxigenic *Escherichia coli* O157:H7 strain NCTC12900 causes attaching-effacing lesions and *eae*-dependent persistence in weaned sheep. Int. J. Med. Microbiol. **293:**299-308.

43. **Woodward, M. J., D. Gavier-Widen, I. M. McLaren, C. Wray, M. Sozmen, and G. R. Pearson.** 1999. Infection of gnotobiotic calves with *Escherichia coli* O157:H7 strain A84. Vet. Rec. **144:**466-470.

CHAPTER 4. EXPLORING THE RESPONSE OF *ESCHERICHIA COLI* O157:H7 EDL933 WITHIN *ACANTHAMOEBA CASTELLANII* BY GENOME-WIDE TRANSCRIPTIONAL PROFILING

A paper to be submitted to Applied and Environmental Microbiology

Michael D. Carruthers, Bryan H. Bellaire and F. Chris Minion*

Department of Veterinary Microbiology and Preventive Medicine, and Interdepartmental Microbiology Program, Iowa State University, Ames, IA 50011
*Corresponding Author, fcminion@iastate.edu

Abstract

Free-living protozoa, such as *Acanthamoeba castellanii*, are environmental hosts for serveral bacterial species pathogenic to humans and animals. Intracellular residence witin protozoa has been implicated in enhancing virulence by selecting for virulence factors and antibiotic resistance outside of the animal host. To better understand this relationship with *E. coli* O157:H7, we characterized *E. coli* O157:H7 transcriptomes from extracellular bacteria in grown in broth to *E. coli* O157:H7 residing within *A. castellanii* using two-color microarrays. Statistical analysis indicated that 969 genes were differentially expressed at a p value < 0.018, with a false discovery rate of 1.9% and a fold change cutoff of 1.3. There were 655 up-regulated transcripts that include 40 genes involved in virulence of which 32 are encoded on O-islands. These up-regulated virulence-associated genes included shiga toxin genes (*stx*1A, *stx*1B *stx*2A), 14 genes involved in LEE and non-LEE encoded type three secretion system components and secreted factors. Additional induced genes include SOS response genes such as *lexA* and *recA*, genes involved in or predicted to be involved in antibiotic resistance (*rarD, macAB,*

marABR, mdtK, yojI, yhgN), the quorum sensing operon *lsrACDB*, the *efe* and *feo* iron acquisition systems There were 314 down-regulated transcripts that include 19 transcripts associated with virulence, 7 of which are encoded on O-islands. Our results demonstrate a significant portion of the *E. coli* O157:H7 genome, including many virulence related genes, was differentially expressed as a result of the intracellular environment.

Introduction

Escherichia coli O157:H7 causes food-borne illness in humans with disease manifested as acute gastroenteritis with symptoms ranging from mild diarrhea to hemorrhagic colitis (39). A potentially fatal sequelae of *E. coli* O157:H7 infection, hemolytic uremic syndrome, is the leading cause of acute renal failure in children (39). Although *E. coli* O157:H7 can be harbored within many wild and domestic animals, healthy cattle are considered the major reservoir (9, 18) with shedding of *E. coli* O157:H7 at levels of 10^2 to 10^5 colony forming units per gram of feces (12, 47). Consequently, the soil and surface water in the vicinity of colonized cattle has been shown to become contaminated with *E. coli* O157:H7 and is the main source of contamination of non-meat food products (33).

Environmental contamination by *E. coli* O157:H7 may occur for a considerable amount of time as it has been shown that theses organisms can persist in manure and slurry (4, 10, 27, 28) and in soil, water, sediment, and animal carcasses (34). While *E. coli* O157:H7 is not thought of as an intracellular pathogen, it has been shown to survive within human macrophage for at least 24 h (42) and the soil protozoan *Acanthamoeba polyphaga* for at least 45 days (7). This bacterial-protozoan interaction has certain implications as protozoa are widely acknowledged as reservoirs for bacterial pathogens such as *Legionella, Listeria*, *Campylobacter, Pseudomonas*, *Helicobacter*, *Mycobacterium*, *Coxellia*, *Salmonella, Staphylococcus,* and harboring of these pathogens within in protozoa has been associated with increased survival and persistence in environment (24), increased virulence (16, 43), and increased resistance to antibiotics (8, 36). With this in mind, protozoa may serve as a vehicle for *E. coli* O157:H7

environmental persistence and transmission as well affecting *E. coli* O157:H7 while it travels through the rumen.

The goal of this study was to better characterize the response of *E. coli* O157:H7 to the intracellular protozoan environment. To accomplish this we employed microarrays to measure the transcriptional changes that occur in *E. coli* O157:H7 as it encounters the intracellular environment of soil protozoan *Acanthamoeba castellanii*, compared to standard planktonic growth conditions. Our results demonstrate a significant portion of the *E. coli* O157:H7 genome, including many virulence related genes, was differentially expressed as a result of the intracellular environment and the results are similar to transcriptome analysis of *E. coli* O157:H7 within macrophages (42).

Materials and Methods

Strains and culture conditions

Escherichia coli O157:H7 EDL933 (ATCC 43895) was grown in Luria-Bertani (LB) broth at 37˚C. *Acanthamoeba castellanii* (ATCC 30010) was grown axenically in ATCC PYG712 broth at 30˚C. Approximately every two weeks *A. castellanii* cultures were passed according to ATCC instructions. Cultures of *E. coli* O157:H7 cultures were grown overnight at 37˚C in LB broth. These cultures were diluted 1:100 in LB broth and incubated with shaking for 2 h. An estimate of *A. castellanii* cell numbers was obtained using a Coulter particle counter prior to experimentation. *Acanthamoeba castellanii* cultures were centrifuged at 100 x *g* for 5 min, resuspended in fresh PYG712 broth and 6 well cell culture plates were seeded with 1 ml of a suspension at 2 x 10^6 cells per ml. After 2 h of incubation, *E. coli* O157:H7 cultures were centrifuged at 5,000 x *g* for 5 min, resuspended in PYG712 broth and added to 6 well cell culture plates at an multiplicity of infection of 1000:1 (bacteria:protozoa). Experimental controls did not contain *A. castellanii* nor were they treated with antibiotic. Cultures were incubated at 30˚C for 30 min, and then gentamicin was added to wells containing *A. castellanii* to a final concentration of 100 µg/µl to eliminate extracellular bacteria. After 2 h of incubation

with gentamicin, *A. castellanii* cultures were centrifuged at 100 x *g* for 5 min to remove planktonic bacteria and resuspended in PYG712 broth containing 25 µg/µl of gentamicin to prevent growth of extracellular bacteria that may be released from dead protozoa. After 2 h, all cultures were centrifuged at 10,000 x *g* for 30 sec, pellets were resuspended in ice cold RNA stop solution (19% ethanol, 0.1% sodium dodecyl sulfate, 1% acidic phenol), and incubated on ice for 30 min. Samples were then centrifuged at 10,000 x *g* for 5 min at 0°C, and the RNA was immediately extracted from the pellets. For determination of intracellular numbers of *E. coli* O157:H7, cultures were generated in the exact same way as cultures used for RNA extraction. At the end of each time point cultures were subjected to 0.1% sodium dodecyl sulfate for 15 min to lyse *A. castellanii.* These samples were serial diluted in phosphate buffered saline, plated on LB agar, incubated at 37˚C for 24 h prior to enumerate. For To, bacteria were enumerated following the initial 30 min incubation and after washing the protozoa twice.

RNA isolation

RNA isolation, DNase treatment, subsequent purification, and determination of the absence of DNA was conducted as previously described (14). RNA integrity and purity (absence of eukaryotic ribosomal peaks) was determined using the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) with all samples measured having a RNA integrity number of 9.0 or higher and void of detectable eukaryotic ribosomal RNA peaks (data not shown).

Labeled cDNA generation and hybridization

Target generation, labeling, reaction clean-up, hybridization, pre- and post-hybridization washes were all conducted as previously described (14).

Microarray experimental design

Eighteen RNA samples, half from cells within *A. castellanii* and half from planktonic cells, were used for the microarray study. A sample from each treatment was randomly paired with sample from another treatment for hybridization on a two-color

microarray substrate for a total of 9 hybridizations. To account for differences in emission intensity of the Cy dyes used (dye bias), dye assignments for control and treated samples were reversed for 4 of the arrays.

Image acquisition and data analysis

Scanning, image segmentation, and normalization were conducted as previously described (41). The normalized values for duplicate spots were averaged within each array to produce one normalized measure of fluorescence for each probe sequence and each of the experimental units. Mixed model analysis was conducted as previously described (14). Cluster of Orthologous Groups of proteins (COGs) information was obtained from NCBI (http://www.ncbi.nlm.nih.gov).

Validation of microarray data

Semi-quantitative real time polymerase chain reaction (qRT-PCR) was preformed on the same samples used for microarray analysis using primer sets for 8 genes (*dnaK, espA, lpfD, macA, ompA, recA, stx*1A, *stx*2A) to confirm significant transcriptional differences due to treatment. The Express One-Step SYBR GreenER kit (Invitrogen) was used for qRT-PCR with the Mx3005P QPCR System (Stratagene, La Jolla, CA) and MxPro 4.1 software. Reaction volume for each well totaled 15 µl and contained 3.69 µl of water, 7.5 µl qPCR Mix, 1.2 µl of each primer at 2.5 µM, 0.03 µl ROX, 1 µl of sample RNA and 0.375 µl SuperScript III. Six biological replicates for each treatment were randomly chosen for qRT-PCR validation and were run in duplicate. In this case biological replication was favored over technical replication. Gene *btuD* was used as a reference gene because it demonstrated no detectable differential expression due to treatment and had a small variance on the microarrays. The method described in Gallup *et al*. was used for primer optimization, detection of inhibition, and troubleshooting of qRT-PCR (19). A 4 point standard curve was constructed with duplicate samples of a collection of all RNAs (Stock 1) and used for the calculation of efficiencies for target genes and the reference gene (19). The ISU equation was used to calculate fold change between treatment and control samples (19), and the students t-test was used to determine

significance of differences. Sample Cts that were greater than 2 standard deviations from the mean were considered outliers and not used in data analysis. qRT-PCR primers used in this study are listed in Table 4.1.

Table 4.1. qRT-PCR primers used to validate microarray results.

Primer	Sequence 5`-3`	Tm˚C
stx1A-F	TAA TGC AAT TCT GGG AAG CGT GGC	60.2
stx1A-R	ATT GTG CGT AAT CCC ACG GAC TCT	60.3
stx2A-F	CGT TCC GGA ATG CAA ATC AGT CGT	59.9
stx2A-R	AAC TGC TCT GGA TGC ATC TCT GGT	60.3
recA-F	ACC ATC TCT ACC GGT TCG CTT TCA	60
recA-R	GCG TGT TCA GCA TCG ATA AAC GCA	60.2
espA-F	CTG CAA ACA GTG AAG GCG GCT ATT	60.1
espA-R	ATT GCA CAT CAG AAC GTG CAC TCG	60.1
ompA-F	TAC CAG TGG ACC AAC AAC ATC GGT	60.3
ompA-R	AAG TGC TTG GTC TGT ACT TCC GGT	60.3
dnaK-F	AGC TGG CAG GAC AAA GTC GAA GTA	60.2
dnaK-R	CAT TCG CGC CAT CAC CTC TTC AAT	59.5
btuD-F	ACC TGA CAC TGC ATC AGC ACG ATA	60.3
btuD-R	TGG TGC TAC GAC CGA GTT TGT CAT	60.3
macA-F	AAC AAG CAC CGA ACA TTC TGA CGC	60.3
macA-R	TGG ATC GCC AAG TAC CGT AAA CCA	60.3
lpfD-F	GTT GCG ACG AAC ATT CGC ATC AGT	60.2
lpfD-R	AGG GAT GAC GGA ATC GCC GAT AAA	60.2

Microarray data accession

The microarray dataset can be accessed from the National Center for Biological Informatics Gene Expression Omnibus using Series accession number GSE16762 (http://www.ncbi.nlm.nih.gov/geo/).

Results

Intracellular survival

Our initial study was to determine the survivability of *E. coli* O157:H7 in *A. castellanii* under the conditions of the microarray study. As shown in Fig. 41, *E. coli* O157:H7 survives for at least 16 h within *A. castellanii*. Initial colony forming units

(CFU) of *E. coli* O157:H7 began at 10^9 and fell 5 logs during the first two hours before leveling off to 10^4.

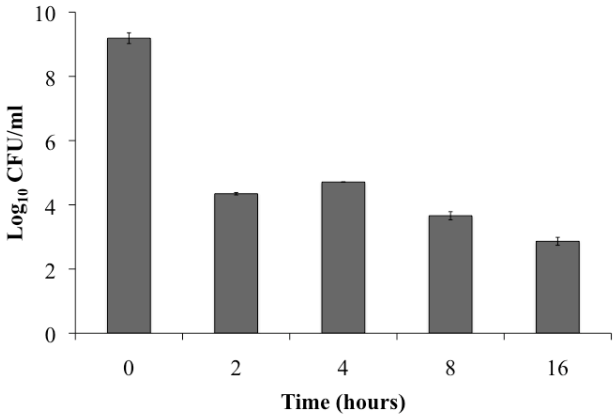

Figure 4.1. Quantification of *E. coli* O157:H7 in *A. castellanii* over time by plate count. Data represents mean ± standard error of three replicates.

Microarrays

To determine the effect of intracellular life on the transcriptome of *E. coli* O157:H7 EDL933, microarrays were used to compare transcript levels of *E. coli* O157:H7 within *A. castellanii* to planktonic cultures. Based on the data from the internal survival curve, an incubation period of 4.5 h was chosen for the microarray study. This included an initial 30 min for *A. castellanii* engulfment of *E. coli*, 2 hours for killing extracellular bacteria, and an additional 2 hours for transcriptional levels to stabilize within viable internal bacteria and for dead bacteria (and its internal RNA) to be degraded. Nine biological replicates were performed and the data from all replicates were used in the statistical analysis.

All RNA preparations were assessed for integrity using the Agilent Bioanalyzer 2100 and for purity using a Nanodrop spectrophotometer. None of the preparations contained measureable *A. castellanii* RNA, even those from mixed bacteria- *A. castellanii* cultures, demonstrating the effectiveness of the bacterial RNA isolation procedure. Also, preliminary studies using 2.5 µg of *A. castellanii*-labeled cDNA hybridized to the *E. coli*

O157:H7 microarray showed minimal reactivity to *E. coli* specific features (data not shown), indicating the specificity of the array. Thus, it is unlikely that protozoa RNA contamination introduced error into the analysis of transcriptional changes in *E. coli* O157:H7.

Statistical analysis indicated that 969 genes with an estimated fold change greater than 1.3 demonstrated transcriptional differences with a p-value < 0.018 and an estimated false discovery rate (FDR) of 1.9%. Significance and differences in transcript levels for all genes are depicted as a volcano plot seen in Fig. 4.2. Of the 969 genes differentially expressed, 655 genes were up-regulated while 314 genes were down-regulated. Differentially expressed genes involved in virulence are listed in Table 4.2. Table 4.3 lists differentially expressed genes associated with antibiotic resistance, the SOS response, and iron acquisition/metabolism. These genes cover 21 clusters of orthologous groups of proteins (COGs) as shown in Fig. 4.3. All statistically significant genes with p <0.05 are listed in supplemental table 4.S1 (Appendix D).

To validate the microarray studies, 8 genes were chosen for qRT-PCR analysis, 6 up-regulated and 2 down-regulated. The gene *btuD* was used for the control since it did not show differential expression in the microarray study. In every case, the qRT-PCR results corroborated the microarray results with respect to direction of differential expression as shown in Fig. 4.4. The degree of transcript difference as measured by qRT-PCR was greater than measured by microarray as has been shown previously (38).

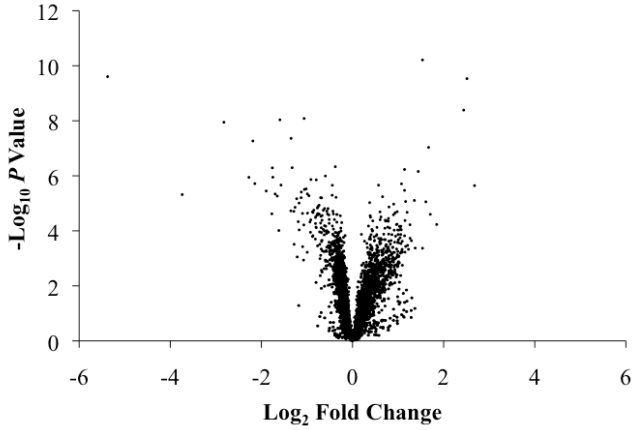

Figure 4.2. Volcano plot of transcriptional differences of *E. coli* O157:H7 within *A. castellanii*. Differences are plotted as Log_2 Fold Change vs.-Log_{10} p-value.

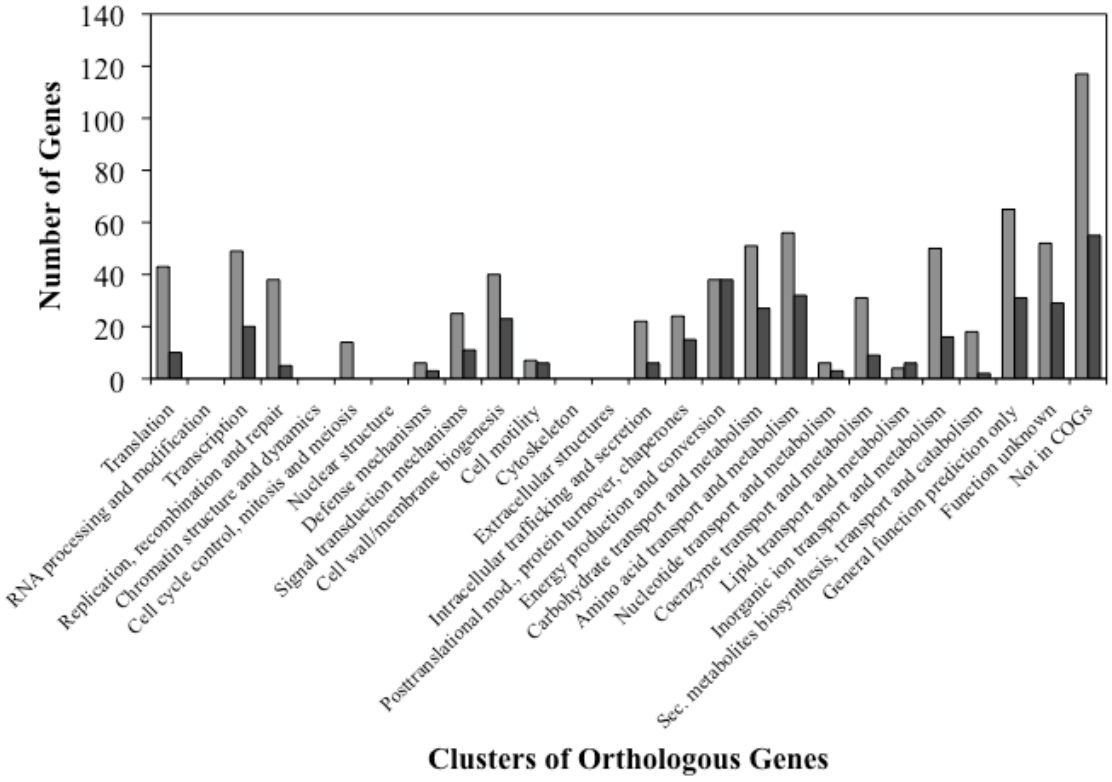

Figure 4.3. Clusters of Orthologous Groups of proteins of the differentially transcribed genes. Light gray bars indicate up-regulated genes; dark gray bars, down-regulated genes.

Table 4.2. Differentially expressed genes involved in virulence as annotated in the ERIC database[1]. All genes were differentially expressed at a p value < 0.04.

Gene	Locus	FC[2]	FC[3]	Function/Product
Up-regulated				
*eivA**	Z4195	1.3		Type III secretion apparatus protein
*eivE**	Z4196	1.3		Putative secreted protein
*eivG**	Z4197	1.3		Type III secretion apparatus protein
*escF**	Z5103	1.4		LEE-encoded type III secretion system component
*espA**	Z5107	1.3	5.0	Secreted protein EspA
*espD**	Z5106	1.3		Secreted protein EspD
*espF**	Z5100	1.3		LEE-encoded type III secreted effector
*espM*1*	Z2565	1.3		Non-LEE-encoded type III secreted effector
*espR*4*	Z3026	1.3		Predicted non-LEE-encoded type III secreted effector
*espY*1*	Z0065	1.4		Non-LEE-encoded type III secreted effector
*espY*2*	Z0078	1.7		Predicted non-LEE-encoded type III secreted effector
*fmlA**	Z2200	1.4		Major subunit of F9 fimbriae
*fmlB**	Z2201	1.3		F9 fimbriae chaperone
*lomP**	Z6028	1.3		Putative Lom-like outer membrane protein of cryptic prophage CP-933P
*lpfD**	Z4966	1.3	2.2	Putative fimbrial protein
*lpfD*2*	Z5221	1.3		Putative fimbrial protein
*lpfE**	Z4965	1.3		Putative fimbrial subunit
lpxC	Z0106	1.5		UDP-3-O-acyl N-acetylglucosamine deacetylase
lpxK	Z1261	1.8		Lipid A 4' kinase
*nleA**	Z6024	1.4		Non-LEE-encoded type III secreted effector
*nleG*2-3*	Z2149	1.3		Predicted non-LEE-encoded type III secreted effector
*nleG*5-2*	Z2151	1.3		Predicted non-LEE-encoded type III secreted effector
orf29*	Z5102	1.3		LEE-encoded predicted type III secretion system factor
*stx*1A*	Z3344	1.3	1.9	Shiga-like toxin 1 subunit A
*stx*1B*	Z3343	1.5	2.2	Shiga-like toxin I subunit B
*stx*2A*	Z1464	1.3		Shiga toxin II subunit A
ycbQ	Z1286	1.8		Predicted fimbrial-like adhesin protein
ydeA	Z2173	1.4		Predicted arabinose transporter
ydeU	Z2196	1.7		C-terminal fragment of a predicted autotransporter
ydhQ	Z2691	1.3		Conserved protein
yfaL	Z3487	1.4		Adhesin
Z0635	Z0635	1.3		Putative membrane spanning export protein

[1] = the ERIC database can be found at http://www.ericbrc.org
[2]FC = microarray estimate of fold change
[3]FC = qRT-PCR estimate of fold change
* indicates that the gene is encoded on an O-island

Table 4.2 (continued)

Gene	Locus	FC[2]	FC[3]	Function/Product
Z1534*	Z1534	1.3		Putative chaperone
Z1544*	Z1544	1.4		Putative acyl-carrier protein synthase
Z1917*	Z1917	1.9		Putative outer membrane protein of prophage CP-933X
Z1964*	Z1964	2.0		Putative iron compound ABC transporter, ATP-binding protein
Z1965*	Z1965	1.9		Putative iron compound ABC transporter/permease
Z2146*	Z2146	1.4		Putative outer membrane protein Lom precursor of prophage CP-933O
Z4383*	Z4383	1.5		Putative iron compound permease protein
Z4385*	Z4385	1.6		Putative ATP-binding protein
Down-regulated				
csgA	Z1676	-1.3		Cryptic curlin major subunit
*epaO**	Z4190	-1.3		Type III secretion apparatus protein
gadE	Z4925	-1.5		Acid-induced positive regulator of glutamate-dependent acid resistance
mviM	Z1705	-1.3		Predicted oxidoreductase
*nleG*2-2	Z2339	-1.4		Non-LEE-encoded type III secreted effector
*nleG*5-1	Z2337	-1.3		Non-LEE-encoded type III secreted effector
*nleH*1-1*	Z0989	-1.3		Non-LEE-encoded type III secreted effector
ompA	Z1307	-2.5	-6.2	Outer membrane protein A
ompX	Z1036	-1.3		Outer membrane protein
*rorf1**	Z5143	-1.3		LEE-encoded type III secretion system factor
sapA	Z2494	-1.3		Predicted antimicrobial peptide transporter subunit
sapF	Z2500	-1.3		Peptide transport system ATP-binding protein
tolC	Z4392	-2.0		Transport channel
toxB	L7095	-1.3		Putative cytotoxin
yehB	Z3277	-1.3		Predicted outer membrane protein
Z0609*	Z0609	-1.4		Predicted protein
Z0615*	Z0615	-1.5		Putative RTX family exoprotein
Z0634*	Z0634	-1.6		Putative cytoplasmic membrane export protein
Z4192*	Z4192	-1.3		Hypothetical protein

Table 4.3. Differentially expressed genes involved in antibiotic resistance, the SOS response and iron acquisition/metabolism. All genes were induced at a p value < 0.022.

Gene	Locus	FC[1]	FC[2]	Function
Antibiotic resistance associated				
fis	Z4621	1.6		Global DNA-binding transcriptional dual regulator
macA	Z1115	1.6	2.1	Macrolide-specific efflux protein
macB	Z1116	1.6		Putative ATP-binding component of a transport system
marA	Z2170	1.4		DNA-binding transcriptional dual activator of multiple antibiotic resistance
marB	Z2169	1.5		Predicted protein
marR	Z2171	1.3		DNA-binding transcriptional repressor of multiple antibiotic resistance
mdtK	Z2690	1.3		Multidrug efflux system transporter
rarD	Z5340	1.3		Predicted chloramphenicol resistance permease
yhgN	Z4798	1.3		Predicted antibiotic transporter
yojI	Z3469	1.7		Fused predicted multidrug transport subunit
SOS response				
dinD	Z5070	1.3		DNA-damage inducible protein
dnaG	Z4419	1.4		DNA primase
ftsK	Z1235	1.7		DNA-binding membrane protein required for chromosome resolution/partitioning
grxA	Z1076	1.7		Glutaredoxin 1, redox coenzyme for ribonucleotide reductase
lexA	Z5642	2.6		DNA-binding transcriptional repressor of SOS regulon
nrdA	Z3489	1.5		Ribonucleoside diphosphate reductase 1 alpha subunit
polB	Z0068	1.5		DNA polymerase II
recA	Z4002	6.4	11.3	DNA strand exchange and recombination
recG	Z5078	1.3		ATP-dependent DNA helicase
recO	Z3846	1.4		Gap repair protein
recX	Z4001	1.4		Inhibitor of RecA
sulA	Z1308	2.2		SOS cell division inhibitor
umuC	Z1947	2.0		DNA polymerase V, subunit C
Iron acquisition/Metabolism				
dps	Z1034	4.6		Fe-binding and storage protein
efeB	Z1521	1.2		Redox component of a tripartite ferrous iron transporter
efeO	Z1520	1.7		Component of a tripartite ferrous iron transporter
efeU	Z1519	1.5		Ferrous iron permease
entA	Z0738	1.7		Enterobactin synthetase component A
entB	Z0737	1.6		Enterobactin synthetase component B
entC	Z0735	1.5		Enterobactin synthetase component C

[1]FC = microarray estimate of fold change
[2]FC = qRT-PCR estimate of fold change

Table 4.3 (continued)

Gene	Locus	FC[1]	FC[2]	Function
entD	Z0723	1.7		Phosphopantetheinyltransferase component of enterobactin synthase multienzyme complex
entE	Z0736	1.8		Enterobactin synthetase component E
entF	Z0727	1.7		Enterobactin synthetase component F
fepA	Z0724	1.8		Iron-enterobactin outer membrane transporter
fepB	Z0734	2.0		Iron-enterobactin transporter subunit
fepC	Z0729	1.8		ATP-binding component of ferric enterobactin transport
fepD	Z0732	1.8		Ferric enterobactin transport
fepE	Z0728	1.5		Ferric enterobactin transport, regulator of length of O-antigen component of LPS chains
fepG	Z0731	1.8		Iron-enterobactin transporter subunit
fes	Z0725	1.6		Enterobactin/ferric enterobactin esterase
Quorum Sensing				
glpD	Z4786	2.1		sn-glycerol-3-phosphate dehydrogenase
lsrA	Z2192	1.3		Fused AI2 transporter
lsrB	Z2189	1.6		AI2 transporter
lsrC	Z2191	1.6		AI2 transporter
lsrD	Z2190	1.7		AI2 transporter
lsrK	Z2194	2.1		Autoinducer-2 kinase

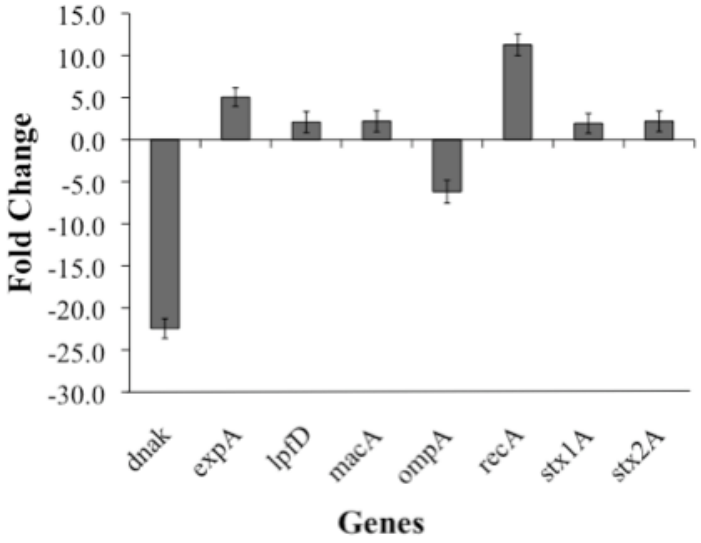

Figure 4.4. Validation of microarray results by qRT-PCR. All results were significant at $p < 0.008$

Discussion

Cattle are the major reservoir of *E. coli* O157:H7 in the United States, colonizing primarily the lower gastrointestinal tract. Outbreaks of foodborne illness have been associated with under cooked contaminated meat products and in other foods through cross contamination. This emphasizes that *E. coli* O157:H7 has adapted to two distinct habitats, the enteric environment of ruminants and humans, and the external environment, namely water, soil and plant surfaces. *Escherichia coli* O157:H7 comes into contact with protozoa while in the rumen and in soil and water that has been contaminated with *E. coli* O157:H7 containing fecal material. Within in the ruminant, *E. coli* O157:H7 must move from the oral cavity down the esophagus where it then must transverse the rumen before ultimately colonizing the colon (40).

The rumen contains a large and diverse population of protozoa that generally require strict anaerobiasis and are bathed in an ill-defined liquid called rumen fluid. Problems arise when attempting to study bacterial-protozoan interactions under natural

conditions. Ruminant protozoa must be maintained under strict anaerobiasis requiring good anaerobic technique, and even under ideal conditions, their life span is short. Rumen fluid contains a plethora of bacteria, which are readily engulfed by the protozoa. Thus, separating the ruminant protozoa from resident bacteria is highly problematic and a potential source of RNA contamination for the microarray studies. A defined system involving a well-studied protozoan such as *A. castellanii* can mitigate many problems of working with ruminant protozoa while allowing the study of bacterial-protozoan interactions. This model can serve a second purpose in studying these types of interactions in the soil and water where *E. coli* O157:H7 can persist and protozoa have been called the "Trojan horses of the microbial world" (6).

There are a limited number of studies involving global analysis of gene expression of bacterial pathogens interacting with host cells. One reason for this is inherent in microarray technology; host transcripts potentially cross-hybridize with organism-specific microarray probes, rendering analysis of data difficult. Eukaryotic biologists circumvent this by using poly-dT primers to specifically prime only polyadenylated mRNA sequences. Bacteriologists study these types of interactions by select capture of transcribed sequences (SCOTS) (20) and by differential lysis. To overcome this problem, we utilized differential lysis in an RNA stabilizing solution. Analysis of the resulting RNA preparations indicated minimal contamination, if any, with *A. castellanii* RNA, simplifying the microarray analysis.

Predation of bacteria by *A. castellanii,* as in the case of *E. coli* K-12, can result in bacterial cell death (2). It is reasonable to think that *E. coli* O157:H7 is exposed to a hostile environment inside *A. castellanii* that would be similar to that found within ruminant protozoa. Supporting this idea, were the up-regulation of transcripts associated with the SOS response within *A. castellanii* (Table 3). The SOS response regulator *lexA* and LexA regulated genes included *recA*, the SOS response repressor gene *lexA,* DNA-damage inducible transcript *dinB,* DNA polymerase V subunit genes *umuCD,* DNA polymerase II transcript *polB*, cell division inhibitor gene *sulA,* ATP-dependent DNA helicase gene *recG*, and gap repair transcript *recO* were all up-regulated on the

microarray. Genes involved in the SOS response but regulated independently of LexA (*recX, nrdA, dnaG*) were also up-regulated. Superoxide dismutase, encoded by *sodC*, was up-regulated indicating that *E. coli* is responding to the production of reactive oxygen species. Similar regulation of the stress and SOS responses has been seen in *E. coli* O157:H7 within human macrophages (42).

In addition to being protected from antibiotics within protozoa (36), pathogenic bacteria isolated from within protozoa have been shown to demonstrate an antibiotic resistant phenotype (8). These previous observations may be explained by the induction of the SOS response in generating random mutations, which have been shown to play a role in *in vivo* antibiotic resistance to ciprofloxacin (17) and survival in the presence of ß-lactams (29, 35). The results of this study show that there is up-regulation of genes associated with increased antibiotic resistance and acquisition of the antibiotic resistant phenotype. Genes up-regulated that are associated with antibiotic resistance include the predicted chloramphenicol permease *rarD*, the macrolide-specific efflux pump *macAB* (26), the *marRAB* operon, which has been shown to be involved in regulating and conferring antibiotic resistance (1), the accessory transcriptional activator of the *mar* operon *fis*, the *mdtK* multidrug efflux transporter, the predicted multidrug transport subunit *yojI*, and *yhgN*, a predicted antibiotic transporter.

Transition metals such as iron are necessary elements for most life, and free iron is limiting *in vivo* (3). The response of *E. coli* to an iron-limiting environment is to synthesize iron-scavenging molecules or siderophores such as enterobactin, also known as enterochelin, in order to acquire iron (3). Genes involved in the biosynthesis of enterobactin (*entABCE, entD, entF)* were up-regulated in this study. Once iron is bound to enterobactin, the iron-enterobactin complex is transported into the cytoplasm. This transport requires the iron-enterobactin transport system encoded by *fepA, fepB, fepCDG, and fepE,* all of which were up-regulated during this study. It is worthy to note that *fepC* along with another up-regulated gene *prrA* are part of the *prrA-modD-yc73-fepC* gene cluster that has been found in 43 strains of *E. coli* O157:H7 including EDL933 but is absent in 19 stains of *E. coli* O157:H7 lacking shiga-toxin encoding genes (49). It has

been suggested that this may be a link between iron metabolism and Stx-mediated virulence. The iron uptake *efeBO* and *efeU* system was also induced. This system has been shown to be activated in a low pH, aerobic environment, exhibit a preference for ferrous iron, and is inactive by mutation in *E. coli* K-12 but active in *E. coli* O157:H7 (13). Once in the cytoplasm, Fe(III) is reduced and removed from enterobactin using the gene product of *fes*, which was also up-regulated in this study. The iron storage protein encoding transcript, *dps*, was down-regulated (-4.6 fold). These results indicate that *E. coli* O157:H7 is selectively regulating genes required for iron assimilation within *A. castellanii*, which is a different set of iron uptake genes from those regulated in response to growth within human macrophages (42). Additionally, this indicates that the environment within *Acanthamoeba* is iron limiting.

Lipopolysaccharides (LPS) and OmpA have been shown to play a crucial role in the interaction of *E. coli* K1 and *A. castellanii* (2). In this study, *ompA* was down-regulated in response to the *A. castellani* intracellular environment and genes involved in LPS synthesis and modification were up-regulated. The *fepE* gene product has been shown to have a second function in *Salmonella*. Besides participating in the enterobactin transport system, FepE participates in producing very long O-antigen that plays a role in complement resistance (11). The O-antigen polymerase encoding transcript *wzy* was up-regulated as well. Thus it appears that *E. coli* O157:H7 may be lengthening the O-antigen portion of LPS while inside *A. castellanii*, which may provide a measure of protection.

Exopolysaccharides produced on the extracellular surface of bacteria have been shown to enhance virulence of bacterial pathogens. *Escherichia coli* O157:H7 produces colanic acid on its surface (23), which has been shown to provide protection from osmotic, oxidative (15), acid, and heat stresses (30). Colanic acid also aids in attachment to alfalfa sprouts (32) and plays a role in survival in simulated gastrointestinal fluids (31). The genes that encode for the latter steps in the synthesis of colanic acid, *wza, wxb,* and *wzb,* were up-regulated in *E. coli* O157:H7 within *A. castellanii*. The gene encoding the WzxE flippase, which mediates movement of the lipid-capsular polysaccharide precursor from the cytoplasmic to periplasmic leaflet of the inner membrane, was also up-regulated.

These results suggest that the concentration of colanic acid on the bacterial surface is increased in the *A. castellanii* environment.

A recent study has implicated autoinducer-2 (AI-2) in the regulation of certain gene virulence genes in *E. coli* O157:H7 (5). Following the addition of AI-2, temporal gene expression changes resulted in the up-regulation of the *lsr* AI-2 uptake operon, flagella/fimbrae genes, iron-related genes, colanic acid synthesis, and LEE encoded genes (5). The AI-2 transporter and kinase-encoding transcript *lsrACDB* responsible for uptake of AI-2 was up-regulated in our study. In addition to AI-2 uptake genes, our study also demonstrated up-regulation of LEE encoded virulence genes, iron acquisition/metabolism genes, up-regulation of certain fimbrae genes (*lpfD, lpfD2, lpfE, ycbQ, ydeA*) and colanic acid biosynthesis genes that was seen previously in the *E. coli* O157:H7 response to AI-2 (5). *glpD*, which has been shown to prevent *lsr* repression by metabolizing glycerol-3-phoshate to prevent built up this compound and thereby indirectly inhibit *lsr* expression (48) was up-regulated. These results together imply that *E. coli* O157:H7 may be involved in AI-2 mediated quorum sensing within *A. castellanii*. Further studies will be needed to confirm this.

The maintenance and expression of bacterial virulence genes involved in human and animal infections have been seen in bacterial-protozoan interactions (37, 43). Multiple transcripts associated with virulence, as annotated in the ERIC database (21), were up-regulated in our microarray study. Shiga toxins have been shown to play a role in survival within grazing protozoa, and it has been postulated that the maintenance of shiga toxin genes is important for the protozoa-bacterial interaction and not the mammalian host interaction (45). Our microarrays reinforce this hypothesis by demonstrating the up-regulation of shiga toxin encoding transcripts *stx*1A, *stx*1B and *stx*2A.

The role of the locus of enterocyte effacement (LEE) encoded type three secretion system (T3SS) in the development of *E. coli* O157:H7 attaching and effacing lesions and translocation of effectors is well documented (25, 44, 46). Our microarray analysis indicated that three genes that comprise part of the LEE encoded T3SS (*espAD, escF*) on

the transcript *expADB-escF-Z*5102-*Z*5104 were up-regulated. This result was further confirmed by qRT-PCR analysis estimating the up-regulation of *espA* to be 5.1 fold.

A second T3SS has been described in *E. coli* O157:H7, which shares homology with the *Salmonella* pathogenicity island 1 system of *Salmonella enterica* (22). This second T3SS is found in many *E. coli* strains in mutated form and has been implicated in the virulence of a septicemic *E. coli* strain (22). Three genes in this T3SS *eivCAEGF* operon, *eivAEG,* were up-regulated as well. Factors secreted by T3SS up-regulated in our study include two LEE encoded and seven non-LEE encoded translocated factors. There are several bacterial pathogens that use a T3SS in both intracellular and extracellular infections. These results suggest there may be a role for T3SS and secreted effectors in interactions between *E. coli* O157:H7 and *A. castellani*.

In summary, the analysis of transcript changes in *E. coli* O157:H7 during its interaction with *A. castellani* show that *E. coli* up-regulates genes involved with managing various stressful environments including limited iron and oxidative stress. The phenotype changes required for protozoa survival may also include changes in the surface architecture and the translocation of effector molecules by T3SS. Some virulence genes were also up-regulated suggesting that protozoa may serve as the environment that selects for and helps to maintain virulence genes that result in colonization and disease outbreaks in mammalian populations.

Acknowledgements

We would like to thank Dr. Greg Phillips for providing *A. castellanii*. This study was funded in part by a contract to FCM from the United States Department of Agriculture (Specific Cooperative Agreement 58-3625-2-127).

References

1. **Alekshun, M. N., and S. B. Levy.** 1997. Regulation of chromosomally mediated multiple antibiotic resistance: the *mar* regulon. Antimicrob. Agents Chemother. **41:**2067-2075.
2. **Alsam, S., S. R. Jeong, J. Sissons, R. Dudley, K. S. Kim, and N. A. Khan.** 2006. *Escherichia coli* interactions with *Acanthamoeba*: a symbiosis with environmental and clinical implications. J. Med. Microbiol. **55:**689-694.
3. **Andrews, S. C., A. K. Robinson, and F. Rodriguez-Quinones.** 2003. Bacterial iron homeostasis. FEMS Microbiol. Rev. **27:**215-237.
4. **Avery, L. M., K. Killham, and D. L. Jones.** 2005. Survival of *E. coli* O157:H7 in organic wastes destined for land application. J. Appl. Microbiol. **98:**814-822.
5. **Bansal, T., P. Jesudhasan, S. Pillai, T. K. Wood, and A. Jayaraman.** 2008. Temporal regulation of enterohemorrhagic *Escherichia coli* virulence mediated by autoinducer-2. Appl. Microbiol. Biotechnol. **78:**811-819.
6. **Barker, J., and M. R. Brown.** 1994. Trojan horses of the microbial world: protozoa and the survival of bacterial pathogens in the environment. Microbiology **140 (Pt 6):**1253-9.
7. **Barker, J., T. J. Humphrey, and M. W. Brown.** 1999. Survival of *Escherichia coli* O157 in a soil protozoan: implications for disease. FEMS Microbiol. Lett. **173:**291-295.
8. **Barker, J., H. Scaife, and M. R. Brown.** 1995. Intraphagocytic growth induces an antibiotic-resistant phenotype of *Legionella pneumophila*. Antimicrob. Agents Chemother. **39:**2684-2688.
9. **Barkocy-Gallagher, G. A., T. M. Arthur, M. Rivera-Betancourt, X. Nou, S. D. Shackelford, T. L. Wheeler, and M. Koohmaraie.** 2004. Characterization of O157:H7 and other *Escherichia coli* isolates recovered from cattle hides, feces, and carcasses. J. Food. Prot. **67:**993-998.

10. **Bolton, D. J., C. M. Byrne, J. J. Sheridan, D. A. McDowell, and I. S. Blair.** 1999. The survival characteristics of a non-toxigenic strain of *Escherichia coli* O157:H7. J. Appl. Microbiol. **86:**407-411.

11. **Bravo, D., C. Silva, J. A. Carter, A. Hoare, S. A. Alvarez, C. J. Blondel, M. Zaldivar, M. A. Valvano, and I. Contreras.** 2008. Growth-phase regulation of lipopolysaccharide O-antigen chain length influences serum resistance in serovars of *Salmonella*. J. Med. Microbiol. **57:**938-946.

12. **Campbell, G. R., J. Prosser, A. Glover, and K. Killham.** 2001. Detection of *Escherichia coli* O157:H7 in soil and water using multiplex PCR. J. Appl. Microbiol. **91:**1004-1010.

13. **Cao, J., M. R. Woodhall, J. Alvarez, M. L. Cartron, and S. C. Andrews.** 2007. EfeUOB (YcdNOB) is a tripartite, acid-induced and CpxAR-regulated, low-pH Fe2+ transporter that is cryptic in *Escherichia coli* K-12 but functional in *E. coli* O157:H7. Mol. Microbiol. **65:**857-875.

14. **Carruthers, M. D., and C. Minion.** 2009. Transcriptome analysis of *Escherichia coli* O157:H7 EDL933 during heat shock. FEMS Microbiol. Lett. **295:**96-102.

15. **Chen, J., S. M. Lee, and Y. Mao.** 2004. Protective effect of exopolysaccharide colanic acid of *Escherichia coli* O157:H7 to osmotic and oxidative stress. Int. J. Food. Microbiol. **93:**281-286.

16. **Cirillo, J. D., S. Falkow, and L. S. Tompkins.** 1994. Growth of *Legionella pneumophila* in *Acanthamoeba castellanii* enhances invasion. Infect. Immun. **62:**3254-3261.

17. **Cirz, R. T., J. K. Chin, D. R. Andes, V. de Crecy-Lagard, W. A. Craig, and F. E. Romesberg.** 2005. Inhibition of mutation and combating the evolution of antibiotic resistance. PLoS Biol. **3:**e176.

18. **Dean-Nystrom E.A, B. B. T., Moon H.W, O'Brien A.D.** 1998. Bovine infection with Shiga-toxin producing *Escherichia coli*, p. 261-267. *In* O. B. A. D. Kaper J.B (ed.), *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli* strains. American Society for Microbiology, Washington, D.C.

19. **Gallup, J. M., and M. R. Ackermann.** 2006. Addressing fluorogenic real-time qPCR inhibition using the novel custom Excel file system 'Focusfield2-6GallupqPCRSet-upTool-001' to attain consistently high fidelity qPCR reactions. Biol. Proceed. Online **8:**87-152.

20. **Graham, J. E., and J. E. Clark-Curtiss.** 1999. Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS). Proc. Natl. Acad. Sci. U S A **96:**11554-11559.

21. **Greene, J. M., G. Plunkett, 3rd, V. Burland, J. Glasner, E. Cabot, B. Anderson, E. Neeno-Eckwall, Y. Qiu, B. Mau, M. Rusch, P. Liss, T. Hampton, D. Pot, M. Shaker, L. Shaull, P. Shetty, C. Shi, J. Whitmore, M. Wong, S. Zaremba, F. R. Blattner, and N. T. Perna.** 2007. A new asset for pathogen informatics--the Enteropathogen Resource Integration Center (ERIC), an NIAID Bioinformatics Resource Center for Biodefense and Emerging/Re-emerging Infectious Disease. Adv. Exp. Med. Biol. **603:**28-42.

22. **Ideses, D., U. Gophna, Y. Paitan, R. R. Chaudhuri, M. J. Pallen, and E. Z. Ron.** 2005. A degenerate type III secretion system from septicemic *Escherichia coli* contributes to pathogenesis. J. Bacteriol. **187:**8164-8171.

23. **Junkins, A. D., and M. P. Doyle.** 1992. Demonstration of exopolysaccharide production by enterohemorrhagic *Escherichia coli*. Curr. Microbiol. **25:**9-17.

24. **King, C. H., E. B. Shotts, Jr., R. E. Wooley, and K. G. Porter.** 1988. Survival of coliforms and bacterial pathogens within protozoa during chlorination. Appl. Environ. Microbiol. **54:**3023-3033.

25. **Knutton, S., I. Rosenshine, M. J. Pallen, I. Nisan, B. C. Neves, C. Bain, C. Wolff, G. Dougan, and G. Frankel.** 1998. A novel EspA-associated surface organelle of enteropathogenic *Escherichia coli* involved in protein translocation into epithelial cells. EMBO J. **17:**2166-2176.

26. **Kobayashi, N., K. Nishino, and A. Yamaguchi.** 2001. Novel macrolide-specific ABC-type efflux transporter in *Escherichia coli*. J. Bacteriol. **183:**5639-5644.

27. **Kudva, I. T., K. Blanch, and C. J. Hovde.** 1998. Analysis of *Escherichia col*i O157:H7 survival in ovine or bovine manure and manure slurry. Appl. Environ. Microbiol. **64:**3166-3174.

28. **Lau, M. M., and S. C. Ingham.** 2001. Survival of faecal indicator bacteria in bovine manure incorporated into soil. Lett. Appl. Microbiol. **33:**131-136.

29. **Levin, B. R.** 2004. Microbiology. Noninherited resistance to antibiotics. Science **305:**1578-1579.

30. **Mao, Y., M. P. Doyle, and J. Chen.** 2001. Insertion mutagenesis of wca reduces acid and heat tolerance of enterohemorrhagic *Escherichia coli* O157:H7. J. Bacteriol. **183:**3811-3815.

31. **Mao, Y., M. P. Doyle, and J. Chen.** 2006. Role of colanic acid exopolysaccharide in the survival of enterohaemorrhagic *Escherichia coli* O157:H7 in simulated gastrointestinal fluids. Lett. Appl. Microbiol. **42:**642-647.

32. **Matthysse, A. G., R. Deora, M. Mishra, and A. G. Torres.** 2008. Polysaccharides cellulose, poly-beta-1,6-n-acetyl-D-glucosamine, and colanic acid are required for optimal binding of *Escherichia coli* O157:H7 strains to alfalfa sprouts and K-12 strains to plastic but not for binding to epithelial cells. Appl. Environ. Microbiol. **74:**2384-2390.

33. **McGee, P., D. J. Bolton, J. J. Sheridan, B. Earley, G. Kelly, and N. Leonard.** 2002. Survival of *Escherichia coli* O157:H7 in farm water: its role as a vector in the transmission of the organism within herds. J. Appl. Microbiol. **93:**706-713.

34. **Mead, P. S., and P. M. Griffin.** 1998. *Escherichia coli* O157:H7. Lancet **352:**1207-1212.

35. **Miller, C., L. E. Thomsen, C. Gaggero, R. Mosseri, H. Ingmer, and S. N. Cohen.** 2004. SOS response induction by beta-lactams and bacterial defense against antibiotic lethality. Science **305:**1629-1631.

36. **Miltner, E. C., and L. E. Bermudez.** 2000. *Mycobacterium avium* grown in *Acanthamoeba castellanii* is protected from the effects of antimicrobials. Antimicrob. Agents Chemother. **44:**1990-1994.

37. **Molmeret, M., M. Horn, M. Wagner, M. Santic, and Y. Abu Kwaik.** 2005. Amoebae as training grounds for intracellular bacterial pathogens. Appl. Environ. Microbiol. **71:**20-88.

38. **Morey, J. S., J. C. Ryan, and F. M. Van Dolah.** 2006. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. Biol. Proced. Online. **8:**175-193.

39. **Nataro, J. P., and J. B. Kaper.** 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11:**142-201.

40. **Naylor, S. W., J. C. Low, T. E. Besser, A. Mahajan, G. J. Gunn, M. C. Pearce, I. J. McKendrick, D. G. Smith, and D. L. Gally.** 2003. Lymphoid follicle-dense mucosa at the terminal rectum is the principal site of colonization of enterohemorrhagic *Escherichia coli* O157:H7 in the bovine host. Infect. Immun. **71:**1505-1512.

41. **Oneal, M. J., E. R. Schafer, M. L. Madsen, and F. C. Minion.** 2008. Global transcriptional analysis of *Mycoplasma hyopneumoniae* following exposure to norepinephrine. Microbiology **154:**2581-2588.

42. **Poirier, K., S. P. Faucher, M. Beland, R. Brousseau, V. Gannon, C. Martin, J. Harel, and F. Daigle.** 2008. *Escherichia coli* O157:H7 survives within human macrophages: global gene expression profile and involvement of the Shiga toxins. Infect. Immun. **76:**4814-4822.

43. **Rasmussen, M. A., S. A. Carlson, S. K. Franklin, Z. P. McCuddin, M. T. Wu, and V. K. Sharma.** 2005. Exposure to rumen protozoa leads to enhancement of pathogenicity of and invasion by multiple-antibiotic-resistant *Salmonella enterica* bearing SGI1. Infect. Immun. **73:**4668-4675.

44. **Roe, A. J., D. E. Hoey, and D. L. Gally.** 2003. Regulation, secretion and activity of type III-secreted proteins of enterohaemorrhagic *Escherichia coli* O157. Biochem. Soc. Trans. **31:**98-103.

45. **Steinberg, K. M., and B. R. Levin.** 2007. Grazing protozoa and the evolution of the *Escherichia coli* O157:H7 Shiga toxin-encoding prophage. Proc. Biol. Sci. **274:**1921-1929.

46. **Tobe, T., S. A. Beatson, H. Taniguchi, H. Abe, C. M. Bailey, A. Fivian, R. Younis, S. Matthews, O. Marches, G. Frankel, T. Hayashi, and M. J. Pallen.** 2006. An extensive repertoire of type III secretion effectors in *Escherichia coli* O157 and the role of lambdoid phages in their dissemination. Proc. Natl. Acad. Sci. U S A **103:**14941-14946.

47. **Wang, G., T. Zhao, and M. P. Doyle.** 1996. Fate of enterohemorrhagic *Escherichia coli* O157:H7 in bovine feces. Appl. Environ. Microbiol. **62:**2567-2570.

48. **Xavier, K. B., and B. L. Bassler.** 2005. Regulation of uptake and processing of the quorum-sensing autoinducer AI-2 in *Escherichia coli*. J. Bacteriol. **187:**238-248.

49. **Ye, C., and J. Xu.** 2001. Prevalence of iron transport gene on pathogenicity-associated island of uropathogenic *Escherichia coli* in *E. coli* O157:H7 containing Shiga toxin gene. J. Clin. Microbiol. **39:**2300-2305.

CHAPTER 5. GENERAL CONCLUSIONS

General Discussion

Escherichia coli O157:H7 continues to be a significant foodborne pathogen of humans. Although several aspects of *E. coli* O157:H7 virulence in humans and colonization of ruminants have been elucidated, there are still many questions left unanswered.

Microbes are subjected to changing environmental conditions that affect many different metabolic pathways. When bacteria are subjected to stress, they differentially express genes to cope with that stress to enhance their survival. Measuring this altering gene expression is one way of characterizing the response to a changing environment. This dissertation research focused on the transcriptional changes that occur in *E. coli* O157:H7 in environments that it encounters within the major reservoir for *E. coli* O157:H7, the ruminant. A custom built PCR-based microarray was used to follow these transcriptional changes.

To determine the experimental conditions that would be examined in these studies, several possibilities were considered. First, a general overview of the environments that *E. coli* O157:H7 would encounter as it traversed through the ruminant gastrointestinal tract was outlined. That list also included conditions associated with food processing. As this list was considered, those that could be mimicked *in vitro* were considered in more detail. Finally, the experiments that offered the greatest likelihood of producing meaningful results were then chosen for further study.

The first study involved heat shock, which served two purposes. First, the custom-built microarray needed to be validated with a study that could be compared to other studies in the literature to corroborate the array results. The study also needed to be unique and additive to the literature. Second, the experiment should mimic the food-processing environment where high temperatures are often used to clean and sanitize equipment to prevent contamination. Thus, a somewhat higher temperature shift of 37°C – 50°C was chosen for analysis, and the more traditional 30°C – 42°C shift was added to

confirm the initial results. Our results demonstrated that heat shock induced marked up-regulation in genes involved in the classical heat shock response of *E. coli* K-12 and *E. coli* O157:H7 specific genes, confirming the array and establishing its usefulness for other studies. The heat shock response is conserved and our transcriptome characterization confirms that this is the case in *E. coli* O157:H7. It is unlikely that *E. coli* O157:H7 specific genes up-regulated during heat shock are involved in the classical heat shock response. Rather, it is thought that heat shock serves as an environmental cue resulting in the induction of genes associated with survival within the host.

The remaining portion of the dissertation research focused on the environments within the ruminant gastrointestinal tract. Once cattle consume *E. coli* O157:H7, it enters the upper gastrointestinal tract traveling through a series of organs with quite different environments and spends a significant amount of time in the rumen and intestine. Common environmental factors in the gastrointestinal tract include a lack of oxygen (anaerobiasis), an increased temperature compared to that of the external environment, and exposure to rumen fluid, at least in the initial stages following ingestion. It is also prudent to point out that the relative lack of oxygen and increased temperature over ambient are both common factors to ruminant colonization and infection in humans. The decision to measure heat shock responses independently of other environmental changes was made based on a long history of heat shock studies in other organisms. The anaerobiasis and rumen fluid treatment studies were combined, however, because they shared one condition (anaerobiasis) and transcriptional changes were similar between the two conditions. Finally, in addition to numerous bacteria, many of which have not been characterized or grown in the laboratory, the rumen contains various species of protozoa that thrive in the anaerobic environment but are difficult to study experimentally outside the cattle rumen. The decision to use *A. castellanii* as a substitute host for *E. coli* O157:H7 was based on its ease of growth in the laboratory and the fact that bacterial carriage in *A. castellanii* has implications in persistence of *E. coli* O157:H7 in the environment.

Following the heat shock studies, a second microarray study was conducted characterizing the transcriptome changes by *E. coli* O157:H7 during growth in a rich medium under normal aerobic conditions (37°C with shaking) compared to anaerobic conditions in the same media at 37°C with shaking. A third array study was then conducted exposing *E. coli* to rumen fluid under anaerobic conditions and comparing that to normal anaerobic conditions in LB medium. These three conditions, temperature shift, anaerobiasis and rumen fluid exposure, represented the non cell-associated environmental variables in the rumen. The results of the microarray analysis indicated that *E. coli* O157:H7 up-regulated genes involved in anaerobic respiration and fermentation along with select virulence factors as a response to limiting oxygen while exposure to rumen fluid under anaerobic conditions had minimal impact on the transcriptome. There were genes commonly regulated between the two conditions, and in almost every instance, rumen fluid enhanced the level of transcriptional changes compared to the anaerobic study. Induction of anaerobic metabolism genes in response to limiting oxygen is obvious, but it is interesting that a select number of virulence associated adherence genes were also induced. This result suggests that lack of oxygen may be an environmental cue for the induction of energy intensive, virulence related processes such as adherence factors in low energy environments.

A final study was conducted to characterize the interaction of *E. coli* O157:H7 with the model protozoan, *A. castellanii*. The results of this analysis touch upon many aspects of intracellular life. Genes involved in iron acquisition and metabolism, the SOS response, antibiotic resistance and virulence genes were all induced as a result of intracellular carriage. The simplest explanation is that the up-regulated genes provide a fitness advantage to *E. coli* O157:H7 in this environment, and thus these genes are more likely to be maintained within the bacterial population. This along with other phenotypes associated with intracellular life, such as acquisition of an antibiotic resistance phenotype, have major implications in clinical human disease. These results underscore the importance of merging environmental microbiology with the study pathogenesis.

In this dissertation, my results show that *E. coli* O157:H7 demonstrates gene regulation similar to that of non-pathogenic *E. coli* and other bacteria in response to the aforementioned environments. In addition, *E. coli* O157:H7 demonstrated differential expression of virulence factors and genes found within other pathogenic *E. coli* along with genes that have not been characterized as being differentially transcribed or associated with the previously described environmental factors. The characterization of *E. coli* O157:H7 within these environments may potentially provide insights into unanswered questions and shine light on yet unknown mechanisms of survival and persistence.

Recommendations for Future Study

Up-regulated *E. coli* O157:H7 genes not found in *E. coli* K-12

The results of this research have laid the foundation for a variety of future studies. In Chapters 2 and 3 we reported that *E. coli* O157:H7 differentially up-regulated genes that are not present in K-12 and that have not previously been associated with heat shock, anaerobiasis and exposure to rumen fluid. I would propose to further characterize the role of these genes in survival during heat shock. Perhaps one could study the regulation of these genes in more detail by identifying regulatory proteins by mutational analysis. Phenotypic changes in these mutants could give insight into their role in survival and pathogenesis. *In silico* analysis followed by *in vitro* studies could help in the identification of conserved motifs up-stream of *E. coli* O157:H7 specific heat shock induced genes as well. Elucidation of the mechanisms by which these genes are regulated may shine light on previously unknown regulatory pathways that may include heat shock as an environmental cue.

Similarly with the studies in Chapter 3, *in silico* analysis followed by *in vitro* studies may help to understand the mechanisms by which *E. coli* O157:H7 specific genes are up-regulated and under what conditions. Transcriptional and translational fusions may be important tools in analyzing these genes. One could generate *lac* fusions in up-regulated genes, and then search for regulators using transposon mutagenesis. A major

finding in Chapter 3 is the induction of curli in response to lack of oxygen and exposure to rumen fluid. Determining the ability of curli mutants to colonize ruminants and the affect on shedding may help to understand why curli are regulated in this way.

Characterization of genes induced as a result of *E. coli* O157:H7 within *Acanthamoeba*

The possibility of future work regarding the research presented in Chapter 4 is, in my opinion, staggering. Not only do the results touch on a number on microbiological disciplines including evolution, pathogenesis and environmental microbiology, but also they highlight many different aspects of *E. coli* O157:H7-protozoan interactions that could be studied.

Iron acquisition/metabolism

Iron acquisition is a major undertaking of any bacterial pathogen as *in vivo* life is inherently iron limiting. Studies involving the specific iron uptake systems that were up-regulated and how they impact the maintenance of *E. coli* O157:H7 within *Acanthamoeba* should be pursued. For instance, one could focus on the iron acquisition system *efeBO/efeU* and its role in pathogenesis. Characterizing mutants in these and other iron acquisition systems for their ability to live within *Acanthamoeba* may shed light on minimal requirements necessary in bacterial-protozoan interactions and why certain systems seem to be favored, at least transcriptionally.

Quorum sensing

Bacterial communication through chemical signals has been shown to be an important way of regulating many genes and phenotypes in different bacteria. The extent to which quorum sensing is important in the intracellular life of *E. coli* O157:H7 has yet to be explored. Data presented in Chapter 4 suggested that *E. coli* O157:H7 are quorum sensing while inside *Acanthamoeba*. I suggest a traditional genetic approach by mutational analysis of the *lsr* operon along with reporter fusions, such as fluorescent proteins, to determine if *E. coli* O157:H7 is quorum sensing within *Acanthamoeba* and what effects do mutations in the quorum sensing system have on the ability of *E. coli*

O157:H7 to be maintained within *Acanthamoeba*. Perhaps quorum sensing has a more important role in intracellular survival than has been previously recognized.

SOS response and antibiotic resistance

Several different systems associated with antibiotic resistances were induced as a result of the intracellular carriage of *E. coli* O157:H7 within *Acanthamoeba*. This may be a mechanism by which antibiotic resistance phenotypes emerge and are maintained. A series of studies on the maintenance of antibiotic resistance plasmids and the evolution of antibiotic resistance markers within *Acanthamoeba* are warranted.

A large set of genes under the umbrella of the SOS regulon were also induced within *Acanthamoeba*. This response is triggered by DNA damage and results in error prone DNA replication. One way in which bacteria acquire antibiotic resistance is through the creation of random mutations. Testing the hypothesis that passage of *E. coli* O157:H7 within protozoa leads to increased antibiotic resistance could be accomplished by serial passage of *E. coli* O157:H7 within *Acanthamoeba* and examining antibiotic resistance profiles against non-passaged wild-type strains.

Role of virulence genes and their maintenance

Several different types of virulence factors were induced as a result of *E. coli* O157:H7-*Acanthamoeba* interactions. Some of these factors have well characterized roles during virulence in humans such as shiga toxins, the LEE-encoded type three secretion system and certain translocated effectors, while others such as specific fimbrae, putative translocated factors, a non-LEE encoded type three secretion system, and lipopolysaccharide do not. Generating mutants and assaying their interaction with *Acanthamoeba* may aid in the characterization of less studied virulence factors.

It has been suggested that protozoa play an important role in the maintenance of virulence factors of various bacterial pathogens. *Escherichia coli* O157:H7 encodes for two sets of shiga toxins that cause significant problems to humans during infection but do not seem to confer any benefits to the bacteria. To explore the possibility of the interaction between *E. coli* O157:H7 and protozoa on the maintenance of bacterial

virulence factors, I propose to serially passage *E. coli* O157:H7 in a rich laboratory medium and in protozoa for a period of time. These cultures can then be assayed for various pathogenic characteristics such as adherence to cell culture, ability to form biofilms, ability to produce toxins, and presence of virulence genes. Passage for hundreds of generations might be necessary to see any effects of the intracellular environment.

APPENDIX A

TABLE 2.S1. DIFFERENTIALLY EXPRESSED GENES OF *E. COLI* O157:H7 DURING HEAT SHOCK

Table 2.S1. Differentially expressed genes of *E. coli* O157:H7 during heat shock.

ID	Locus	Function/Product	*p* value	q value	FC
Down-regulated					
ansB	Z4302	Periplasmic L-asparaginase II	0.0008	0.045	-2.5
aphA	Z5654	Diadenosine tetraphosphatase	0.0034	0.086	-1.2
aspA	Z5744	Aspartate ammonia-lyase aspartase	0.0000	0.002	-2.7
cbpA	Z1418	Curved DNA-binding protein functions closely related to DnaJ	0.0027	0.079	-1.2
crp	Z4718	Cyclic AMP receptor protein	0.0000	0.008	-1.4
crr	Z3683	Glucose-specific PTS system enzyme IIA component	0.0030	0.081	-1.3
csgA	Z1676	Curlin major subunit coiled surface structures cryptic	0.0000	0.006	-2.7
csgB	Z1675	Minor curlin subunit precursor similar to CsgA	0.0001	0.010	-1.7
csgC	Z1677	Putative curli production protein	0.0001	0.010	-1.4
csgD	Z1673	Putative 2-component transcriptional regulator for 2nd curli operon	0.0027	0.079	-1.3
csgE	Z1672	Curli production assembly transport component 2nd curli operon	0.0015	0.066	-1.3
csgF	Z1671	Curli production assembly transport component 2nd curli operon	0.0028	0.079	-1.5
csgG	Z1670	Curli production assembly transport component 2nd curli operon	0.0009	0.048	-1.6
cydA	Z0900	Cytochrome d terminal oxidase polypeptide subunit I	0.0004	0.027	-1.5
cydB	Z0901	Cytochrome d terminal oxidase polypeptide subunit II	0.0000	0.005	-1.7
dctA	Z4942	C4-dicarboxylate transport protein	0.0001	0.016	-1.4
dps	Z1034	DNA protection during starvation conditions	0.0038	0.091	-1.6
fba	Z4263	Fructose-bisphosphate aldolase	0.0000	0.004	-2.0
fdoG	Z5436	Formate dehydrogenase-O, major subunit	0.0017	0.067	-1.8
frdA	Z5762	Fumarate reductase	0.0002	0.023	-2.0
frdB	Z5760	Succinate dehydrogenase	0.0030	0.081	-1.4
frdD	Z5758	Fumarate reductase subunit D	0.0016	0.066	-1.8
fumB	Z5724	Fumarase B fumarate hydratase Class I anaerobic isozyme	0.0028	0.079	-1.5
gadA	Z4930	Glutamate decarboxylase isozyme	0.0003	0.023	-4.0

Table 2.S1. (continued)

ID	Locus	Function/Product	*p* value	q value	FC
galM	Z0926	Galactose-1-epimerase mutarotase	0.0008	0.046	-1.3
gcvH	Z4241	Glycine cleavage system protein H	0.0031	0.081	-1.5
grxB	Z1701	Glutaredoxin 2	0.0009	0.048	-1.7
grxC	Z5037	Glutaredoxin 3	0.0008	0.045	-1.3
hdeA	Z4922	Hypothetical protein	0.0001	0.016	-4.7
hns	Z2013	DNA-binding protein HLP-II HU BH2 HD NS pleiotropic regulator	0.0023	0.078	-1.2
hupA	Z5576	DNA-binding protein HU-alpha HU-2	0.0003	0.025	-1.4
hyaB	Z1390	Hydrogenase-1 large subunit	0.0012	0.058	-1.2
hybA	Z4350	Hydrogenase-2 small subunit	0.0043	0.097	-1.2
icdA	Z1865	Isocitrate dehydrogenase	0.0024	0.078	-1.2
kbl	Z5044	2-amino-3-ketobutyrate coenzyme A ligase	0.0007	0.042	-1.4
lamB	Z5634	Maltoporin precursor	0.0006	0.038	-3.6
malE	Z5632	Periplasmic maltose-binding protein	0.0003	0.024	-3.8
malF	Z5631	Part of maltose permease	0.0001	0.009	-10.9
malK	Z5633	ATP-binding component of transport system for maltose	0.0032	0.084	-2.6
mglB	Z3405	Galactose-binding transport protein receptor for galactose taxis	0.0024	0.078	-2.3
msyB	Z1686	Suppresses mutants lacking function of protein export	0.0037	0.090	-1.4
nrfB	Z5670	Formate-dependent nitrite reductase a penta-haeme cytochrome c	0.0041	0.096	-1.2
nrfC	Z5671	Formate-dependent nitrite reductase Fe-S centers	0.0008	0.046	-1.2
nuoC	Z3545	NADH dehydrogenase I chain C D	0.0032	0.083	-1.3
nuoE	Z3544	ATP synthase subunit E	0.0012	0.058	-1.6
nuoH	Z3541	NADH dehydrogenase subunit H	0.0022	0.075	-1.5
nuoJ	Z3539	NADH dehydrogenase subunit J	0.0016	0.067	-1.4
nuol	Z3540	NADH dehydrogenase subunit I	0.0019	0.072	-1.5
nuoN	Z3534	NADH dehydrogenase subunit N	0.0040	0.094	-1.3

Table 2.S1. (continued)

ID	Locus	Function/Product	*p* value	q value	FC
ompR	Z4760	Osmolarity response regulator	0.0000	0.008	-1.3
pckA	Z4758	Phosphoenolpyruvate carboxykinase	0.0002	0.020	-2.4
pepD	Z0298	Aminoacyl-histidine dipeptidase peptidase D	0.0001	0.012	-1.3
pepT	Z1832	Peptidase T	0.0015	0.066	-1.1
pgk	Z4265	Phosphoglycerate kinase	0.0026	0.079	-1.3
ppsA	Z2731	Phosphoenolpyruvate synthase	0.0015	0.066	-1.4
putA	Z1513	Proline dehydrogenase P5C dehydrogenase	0.0016	0.067	-1.4
rfaF	Z5047	ADP-heptose-LPS heptosyltransferase II LPS core biosynthesis	0.0027	0.079	-1.2
ribB	Z4399	3 4-dihydroxy-2-butanone 4-phosphate synthase	0.0025	0.079	-2.1
sdhA	Z0877	Succinate dehydrogenase catalytic subunit	0.0024	0.078	-1.6
sdhC	Z0875	Succinate dehydrogenase cytochrome b556 subunit	0.0006	0.039	-1.4
slp	Z4908	Outer membrane protein induced after carbon starvation	0.0017	0.067	-1.3
sucC	Z0882	Succinyl-CoA synthetase subunit beta	0.0025	0.079	-1.2
talB	Z0008	Transaldolase	0.0011	0.053	-1.1
tdcA	Z4470	Transcriptional activator of tdc operon	0.0034	0.086	-1.7
tdcB	Z4469	Threonine dehydratase	0.0000	0.005	-3.1
tdcC	Z4468	Anaerobically inducible L-threonine L-serine permease	0.0020	0.073	-1.8
tdcE	Z4466	Probable formate acetyltransferase 3	0.0008	0.046	-1.8
tdH	Z5043	L-threonine 3-dehydrogenase	0.0001	0.010	-1.8
tesA	Z0647	Acyl-CoA thioesterase I also functions as protease I	0.0022	0.076	-1.3
tktA	Z4279	Transketolase	0.0030	0.081	-1.5
tnaA	Z5203	Tryptophanase	0.0022	0.075	-2.1
tnaB	Z5204	Low affinity tryptophan permease	0.0011	0.053	-1.6
tolC	Z4392	Outer membrane channel precursor protein	0.0003	0.027	-1.2
ushA	Z0599	UDP-sugar hydrolase 5 -nucleotidase	0.0033	0.085	-1.1
yaeE	Z0210	Putative transport system permease protein	0.0012	0.057	-1.4

Table 2.S1. (continued)

ID	Locus	Function/Product	*p* value	q value	FC
ybaS	Z0606	Glutaminase	0.0020	0.073	-1.3
ycfP	Z1747	Hypothetical protein	0.0020	0.072	-1.1
yciD	Z2034	Putative outer membrane protein	0.0000	0.009	-1.7
ydaA	Z2435	Hypothetical protein	0.0000	0.006	-2.3
ydiA	Z2732	Hypothetical protein	0.0000	0.006	-1.2
yeaG	Z2823	Hypothetical protein	0.0030	0.081	-1.4
yeaX	Z2846	Putative diogenase beta subunit	0.0036	0.088	-1.1
yeiA	Z3402	Dihydropyrimidine dehydrogenase	0.0001	0.010	-1.7
yfcA	Z3590	Putative structural protein	0.0029	0.081	-1.1
yggL	Z4304	Hypothetical protein	0.0039	0.092	-1.1
ygiB	Z4394	Hypothetical protein	0.0013	0.061	-1.2
yhaR	Z4465	Hypothetical protein	0.0033	0.085	-1.9
yhcB	Z4592	Hypothetical protein	0.0021	0.074	-1.2
yhhX	Z4808	Putative regulator	0.0004	0.027	-1.2
yhiE	Z4925	Hypothetical protein	0.0005	0.035	-1.3
yhiM	Z4890	Hypothetical protein	0.0025	0.079	-1.6
yhiP	Z4896	Putative transport protein	0.0020	0.072	-1.7
ynhD	Z2710	Putative ATP-binding component of a transport system	0.0006	0.037	-1.2
ynhE	Z2711	Hypothetical protein	0.0023	0.076	-1.5
yqhE	Z4365	Putative enzyme	0.0028	0.079	-1.4
Z0968	Z0968	Unknown protein encoded by prophage CP-933K	0.0000	0.003	-2.9
Z1479	Z1479	Unknown protein encoded by bacteriophage BP-933W	0.0001	0.016	-1.1
Z1495	Z1495	Unknown protein encoded by bacteriophage BP-933W	0.0010	0.052	-1.1
Z2211	Z2211	Putative enzyme	0.0015	0.066	-1.2
Z2220	Z2220	Putative enzyme	0.0001	0.010	-1.7
Z2323	Z2323	Hypothetical protein	0.0012	0.057	-1.3

Table 2.S1. (continued)

ID	Locus	Function/Product	*p* value	q value	FC
Z2400	Z2400	Partial putative repressor encoded within prophage CP-933R	0.0027	0.079	-1.3
Z2565	Z2565	Putative chaperone protein	0.0001	0.016	-2.0
Z2869	Z2869	Hypothetical protein	0.0000	0.003	-1.4
Z3132	Z3132	Hypothetical protein	0.0002	0.022	-1.4
Z3332	Z3332	Unknown protein encoded within prophage CP-933V	0.0000	0.006	-3.3
Z3341	Z3341	Unknown protein encoded within prophage CP-933V	0.0006	0.039	-1.6
Z4333	Z4333	Putative cytotoxin	0.0016	0.067	-1.1
Z4351	Z4351	Putative hydrogenase subunit	0.0031	0.082	-1.6
Z4464	Z4464	Putative L-serine deaminase	0.0005	0.035	-1.9
Z4695	Z4695	Putative iron storage homoprotein bacterioferritin	0.0020	0.072	-1.2
Up-regulated					
bioB	Z0994	Biotin synthesis sulfur insertion 3F	0.0004	0.030	1.1
caiE	Z0041	Possible synthesis of cofactor for carnitine racemase & dehydratase	0.0024	0.078	1.0
cca	Z4409	tRNA nucleotidyl transferase	0.0028	0.079	1.1
clpX	Z0543	ATP-dependent protease ATP-binding subunit	0.0043	0.097	1.2
dinI	Z1698	Putative damage induced protein I	0.0000	0.002	1.4
dnaJ	Z0015	Chaperone with DnaK heat shock protein	0.0000	0.001	4.7
dnaK	Z0014	Molecular chaperone DnaK	0.0000	0.001	5.4
fdnI	Z2234	Formate dehydrogenase-N nitrate-inducible cytochrome B556 Fdn gamma subunit	0.0008	0.046	1.1
ftsJ	Z4541	Cell division protein	0.0001	0.014	2.9
gltS	Z5081	Glutamate transport	0.0016	0.067	1.1
grpE	Z3907	Heat shock protein repair	0.0000	0.002	2.7
hflB	Z4540	Degrades sigma32 integral membrane peptidase cell division protein	0.0002	0.022	2.2
hflX	Z5780	Repressor	0.0037	0.090	1.8

Table 2.S1. (continued)

ID	Locus	Function/Product	*p* value	q value	FC
hopD	Z4693	Putative leader peptidase	0.0038	0.091	1.1
hslV	Z5479	ATP-dependent protease peptidase subunit	0.0014	0.062	2.1
htpG	Z0590	Heat shock protein 90	0.0000	0.005	1.9
htpX	Z2876	Heat shock protein HtpX	0.0000	0.000	4.9
hydN	Z4021	Electron transport from formate to hydrogen Fe-S centers	0.0002	0.018	1.2
ibpA	Z5183	Heat shock protein	0.0000	0.002	5.8
ibpB	Z5182	Heat shock protein	0.0000	0.006	2.6
ileS	Z0030	Isoleucyl-tRNA synthetase	0.0037	0.091	1.1
lspA	Z0031	Signal peptidase II	0.0029	0.081	1.2
lytB	Z0034	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	0.0021	0.074	1.1
miaA	Z5778	tRNA delta 2 -isopentenylpyrophosphate transferase	0.0004	0.027	2.7
GroEL	Z5748	Chaperonin GroEL	0.0000	0.001	6.6
GroES	Z5747	Co-chaperonin GroES	0.0000	0.000	7.9
murI	Z5528	Glutamate racemase	0.0040	0.094	2.2
narP	Z3450	Nitrate nitrite response regulator sensor NarQ	0.0018	0.069	1.1
nlpA	Z5147	Lipoprotein-28	0.0006	0.039	1.1
prlC	Z4898	Oligopeptidase A	0.0003	0.025	1.3
rna	Z0755	RNase I	0.0032	0.084	1.1
rpoD	Z4420	RNA polymerase sigma factor	0.0025	0.079	1.4
rpoH	Z4835	RNA polymerase sigma factor	0.0003	0.027	2.5
secG	Z4537	Protein-export membrane protein	0.0003	0.025	2.0
sepZ	Z5122	Putative transport protein	0.0042	0.097	1.1
slpA	Z0033	Probable FKBX-type 16KD peptidyl-prolyl cis-trans isomerase	0.0002	0.022	1.2
topA	Z2536	DNA topoisomerase I	0.0001	0.010	1.4
uxuA	Z5920	Mannonate dehydratase	0.0006	0.037	1.2
yaal	Z0013	Hypothetical protein	0.0005	0.035	1.1

Table 2.S1. (continued)

ID	Locus	Function/Product	*p* value	q value	FC
yaeJ	Z0203	Hypothetical protein	0.0037	0.090	1.1
yafD	Z0232	Hypothetical protein	0.0005	0.035	1.5
yahA	Z0403	Hypothetical protein	0.0018	0.068	1.2
ybaP	Z0601	Putative ligase	0.0001	0.016	1.2
ybbN	Z0645	Putative thioredoxin-like protein	0.0030	0.081	1.2
ybil	Z1024	Hypothetical protein	0.0015	0.066	1.5
yccV	Z1318	Hypothetical protein	0.0027	0.079	1.2
yceP	Z1697	Hypothetical protein	0.0000	0.001	3.5
ycfX	Z1760	Putative NAGC-like transcriptional regulator	0.0013	0.061	1.1
yebL	Z2909	Putative adhesin	0.0025	0.079	1.2
yecG	Z2948	Putative regulator	0.0008	0.045	1.1
yedW	Z3061	Putative 2-component transcriptional regulator	0.0006	0.039	1.1
yfhF	Z3795	Putative regulator	0.0035	0.086	1.1
ygiM	Z4408	Hypothetical protein	0.0001	0.016	1.1
ygjI	Z4431	Putative oxidoreductase	0.0022	0.075	1.1
yhaV	Z4482	Hypothetical protein	0.0013	0.060	1.3
yhbE	Z4546	Hypothetical protein	0.0035	0.086	1.3
yheL	Z4701	Hypothetical protein	0.0009	0.048	1.5
yhgI	Z4769	Hypothetical protein	0.0022	0.075	1.5
ynbE	Z2327	Hypothetical protein	0.0028	0.079	1.1
yrbD	Z4556	Hypothetical protein	0.0005	0.035	1.2
yrfH	Z4754	Hypothetical protein	0.0024	0.078	1.5
Z0233	Z0233	Hypothetical protein	0.0030	0.081	1.4
Z0402	Z0402	Putative beta-barrel outer membrane protein	0.0015	0.066	1.2
Z1059	Z1059	Hypothetical protein	0.0011	0.053	1.6
Z1338	Z1338	DNA replication protein DnaC	0.0038	0.091	1.1

Table 2.S1. (continued)

ID	Locus	Function/Product	*p* value	q value	FC
Z1340	Z1340	Unknown protein encoded by cryptic prophage CP-933M	0.0028	0.079	1.1
Z1443	Z1443	Unknown protein encoded by bacteriophage BP-933W	0.0003	0.025	1.4
Z1453	Z1453	Unknown protein encoded by bacteriophage BP-933W	0.0018	0.068	1.3
Z1456	Z1456	Unknown protein encoded by bacteriophage BP-933W	0.0039	0.092	1.4
Z1648	Z1648	Unknown in putative ISEc8	0.0009	0.048	1.2
Z2394	Z2394	Unknown protein encoded within prophage CP-933R	0.0028	0.079	1.2
Z2585	Z2585	Putative dithiobiotin synthetase	0.0034	0.086	1.1
Z3065	Z3065	Hypothetical protein	0.0026	0.079	1.3
Z3970	Z3970	Hypothetical protein	0.0004	0.027	2.1
Z3982	Z3982	Putative transport protein	0.0019	0.072	1.2
Z3983	Z3983	Hypothetical protein	0.0026	0.079	1.1
Z4321	Z4321	Putative PagC-like membrane protein	0.0002	0.020	1.2
Z4650	Z4650	Putative transferase	0.0005	0.035	1.2
Z4883	Z4883	Hypothetical protein	0.0017	0.067	1.2
Z5121	Z5121	Hypothetical protein	0.0017	0.067	1.1
Z5178	Z5178	Putative PTS component	0.0017	0.067	1.1
Z5745	Z5745	Hypothetical protein	0.0001	0.010	1.6

APPENDIX B

TABLE 3.S1. DIFFERENTIALLY EXPRESSED GENES OF *E. COLI* O157:H7 DURING ANAEROBIC GROWTH

Table 3.S1. Differentially expressed genes of *E. coli* O157:H7 during anaerobic growth

Gene	Locus	Location	*p* value	q value	FC	Function/Product
Down-Regulated						
aceF	Z0125	Backbone	1.1E-02	0.092	-1.0	Dihydrolipoamide acetyltransferase
agaZ	Z4484	Backbone	4.6E-03	0.056	-1.0	Putative tagatose 6-phosphate kinase 2
atpD	Z5230	Backbone	3.6E-03	0.047	-1.1	ATP synthase subunit B
atpG	Z5231	Backbone	8.2E-03	0.078	-1.1	ATP synthase subunit C
borW	Z1474	O-Island	1.3E-02	0.096	-1.1	Putative Bor protein precursor of bacteriophage BP-933W
celD	Z2765	Backbone	1.7E-03	0.031	-1.1	Negative transcriptional regulator of cel operon
cheY	Z2936	Backbone	1.1E-02	0.089	-1.1	Chemotaxis regulator
cpsG	Z3212	Backbone	1.6E-03	0.030	-1.1	Phosphomannomutase
cyoA	Z0535	Backbone	2.6E-04	0.011	-1.1	Cytochrome o ubiquinol oxidase subunit II
cyoB	Z0534	Backbone	6.6E-04	0.019	-1.8	Cytochrome o ubiquinol oxidase subunit I
cyoC	Z0533	Backbone	1.5E-03	0.029	-1.4	Cytochrome o ubiquinol oxidase subunit III
cyoD	Z0532	Backbone	1.1E-03	0.025	-1.7	Cytochrome o ubiquinol oxidase subunit IV
cyoE	Z0531	Backbone	2.1E-04	0.010	-1.1	Protoheme IX farnesyltransferase
cysG	Z4729	Backbone	5.9E-04	0.018	-1.1	Uroporphyrinogen III methylase sirohaeme biosynthesis
fdoG	Z5436	Backbone	7.2E-05	0.006	-1.7	Formate dehydrogenase-O, major subunit
fdoI	Z5434	Backbone	1.4E-03	0.028	-1.2	Formate dehydrogenase cytochrome B556 FDO subunit
fliC	Z3013	Junction	2.2E-04	0.010	-1.6	Flagellar biosynthesis protein
fliZ	Z3011	Backbone	9.4E-03	0.084	-1.1	Hypothetical protein
fusA	Z4698	Backbone	6.6E-06	0.003	-2.4	Elongation factor EF-2
glnA	Z5406	Backbone	7.6E-03	0.074	-1.1	Glutamine synthetase
glnH	Z1033	Backbone	1.1E-02	0.091	-1.2	Glutamine ABC transporter periplasmic-binding protein
glpD	Z4786	Backbone	2.1E-07	0.001	-2.5	sn-glycerol-3-phosphate dehydrogenase aerobic
glpE	Z4785	Backbone	3.6E-03	0.046	-1.1	Thiosulfate sulfurtransferase
glpF	Z5472	Backbone	6.0E-03	0.066	-1.1	Facilitated diffusion of glycerol
hrpB	Z0159	Backbone	9.2E-03	0.083	-1.1	Helicase ATP-dependent

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
infB	Z4529	Backbone	3.6E-04	0.014	-1.2	Translation initiation factor IF-2
kgd	Z0880	Backbone	1.2E-03	0.027	-1.4	2-oxoglutarate dehydrogenase decarboxylase component
L7011	L7011	pO157	6.6E-03	0.069	-1.1	Hypothetical protein
L7056	L7056	pO157	1.3E-02	0.096	-1.0	Replication protein
lldD	Z5032	Backbone	2.7E-03	0.041	-1.3	L-lactate dehydrogenase
manC	Z3195	O-Island	1.3E-02	0.096	-1.2	Mannose-1-P guanosyltransferase
mglC	Z3403	Backbone	1.2E-03	0.026	-1.5	Beta-methylgalactoside transporter inner membrane component
moeA	Z1050	Backbone	1.7E-03	0.031	-1.1	Molybdopterin biosynthesis
motA	Z2944	Backbone	2.3E-03	0.038	-1.0	Flagellar motor protein
ndh	Z1748	Backbone	1.1E-02	0.092	-1.1	Respiratory NADH dehydrogenase
nrdI	Z3976	Backbone	7.6E-03	0.074	-1.0	Hypothetical protein
nuoC	Z3545	Backbone	1.6E-03	0.030	-1.3	NADH dehydrogenase I chain C D
nuoE	Z3544	Backbone	2.2E-03	0.038	-1.5	ATP synthase subunit E
nuoF	Z3543	Backbone	7.2E-03	0.072	-1.1	NADH dehydrogenase I chain F
nuoH	Z3541	Backbone	1.5E-04	0.008	-1.4	NADH dehydrogenase subunit H
nuoJ	Z3539	Backbone	6.5E-04	0.019	-1.3	NADH dehydrogenase subunit J
nuoK	Z3538	Backbone	2.9E-03	0.043	-1.2	NADH dehydrogenase kappa subunit
nuoI		Backbone	6.6E-04	0.019	-1.4	NADH dehydrogenase subunit L
nuoN	Z3534	Backbone	2.0E-03	0.034	-1.2	NADH dehydrogenase subunit N
nusA	Z4530	Backbone	3.1E-03	0.043	-1.1	Transcription elongation factor NusA
obgE	Z4545	Backbone	5.0E-03	0.058	-1.1	Putative GTP-binding factor
oxyR	Z5519	Backbone	8.7E-03	0.081	-1.1	Activator hydrogen peroxide-inducible genes
pepB	Z3790	Backbone	4.8E-03	0.057	-1.0	Aminopeptidase B
putP	Z1515	Backbone	1.1E-02	0.089	-1.1	Major sodium proline symporter
rimM	Z3902	Backbone	1.9E-04	0.009	-1.2	16S rRNA-processing protein
rplB	Z4688	Backbone	1.8E-04	0.009	-1.4	50S ribosomal protein L2

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
rplC	Z4691	Backbone	1.1E-05	0.003	-1.8	50S ribosomal protein L3
rplD	Z4690	Backbone	1.6E-05	0.003	-2.1	50S ribosomal protein L4
rplE	Z4678	Backbone	4.6E-05	0.004	-1.4	50S ribosomal protein L5
rplF	Z4675	Backbone	7.1E-04	0.019	-1.2	50S ribosomal protein L6
rplI	Z5812	Backbone	1.3E-02	0.096	-1.1	50S ribosomal protein L9
rplM		Backbone	1.2E-02	0.092	-1.1	50S ribosomal protein L36
rplN	Z4680	Backbone	6.7E-04	0.019	-1.1	50S ribosomal protein L14
rplO	Z4671	Backbone	1.1E-05	0.003	-1.2	50S ribosomal protein L15
rplP	Z4684	Backbone	1.2E-05	0.003	-1.8	50S ribosomal protein L16
rplQ	Z4664	Backbone	9.2E-03	0.083	-1.1	50S ribosomal protein L17
rplR	Z4674	Backbone	1.3E-03	0.027	-1.3	50S ribosomal protein L18
rplS	Z3900	Backbone	3.2E-03	0.043	-1.1	50S ribosomal protein L19
rplV	Z4686	Backbone	1.3E-05	0.003	-2.0	50S ribosomal protein L22
rplW	Z4689	Backbone	7.1E-03	0.072	-1.1	50S ribosomal protein L23
rplX	Z4679	Backbone	1.7E-04	0.009	-1.5	50S ribosomal protein L24
rpmC	Z4683	Backbone	9.8E-07	0.001	-1.9	50S ribosomal protein L29
rpsC	Z4685	Backbone	5.1E-05	0.004	-1.3	30S ribosomal protein S3
rpsD	Z4666	Backbone	8.6E-03	0.080	-1.3	30S ribosomal protein S4
rpsE	Z4673	Backbone	3.4E-05	0.004	-2.2	30S ribosomal protein S5
rpsG	Z4699	Backbone	1.5E-05	0.003	-2.3	30S ribosomal protein S7
rpsH	Z4676	Backbone	2.9E-05	0.003	-1.6	30S ribosomal protein S8
rpsI	Z4588	Backbone	4.6E-04	0.016	-1.6	30S ribosomal protein S9
rpsJ	Z4692	Backbone	1.9E-05	0.003	-1.6	30S ribosomal protein S10
rpsK	Z4667	Backbone	1.9E-05	0.003	-1.8	30S ribosomal protein S11
rpsM	Z4668	Backbone	2.6E-03	0.041	-1.2	30S ribosomal protein S13
rpsQ	Z4681	Backbone	7.7E-04	0.020	-1.2	30S ribosomal protein S17

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
rpsS	Z4687	Backbone	1.8E-05	0.003	-1.9	30S ribosomal protein S19
sdaB	Z4114	Backbone	3.9E-03	0.050	-1.2	L-serine dehydratase deaminase L-SD2
sdaC	Z4113	Backbone	4.1E-05	0.004	-1.4	Probable serine transporter
sdhA	Z0877	Backbone	7.8E-04	0.020	-1.5	Succinate dehydrogenase catalytic subunit
sdhB	Z0878	Backbone	1.7E-03	0.031	-1.4	Succinate dehydrogenase catalytic subunit
sdhC	Z0875	Backbone	4.1E-04	0.015	-1.2	Succinate dehydrogenase cytochrome b556 subunit
secY	Z4670	Backbone	1.3E-05	0.003	-1.7	Preprotein translocase SecY
sgaU	Z5806	Backbone	2.8E-03	0.042	-1.1	Putative hexulose-6-phosphate isomerase
sodA	Z5453	Backbone	1.2E-02	0.095	-1.6	Superoxide dismutase manganese
speB	Z4281	Backbone	7.0E-04	0.019	-1.1	Agmatinase
srlD	Z4012	Backbone	6.4E-03	0.068	-1.2	Sorbitol-6-phosphate 2-dehydrogenase
sucC	Z0882	Backbone	1.2E-03	0.027	-1.2	Succinyl-CoA synthetase subunit beta
tatC	Z5360	Backbone	1.1E-02	0.092	-1.0	Putative integral membrane protein
tesA	Z0647	Backbone	1.1E-02	0.090	-1.1	Acyl-CoA thioesterase I also functions as protease I
torA	Z1415	Backbone	1.2E-02	0.095	-1.1	Trimethylamine N-oxide reductase subunit
trmD	Z3901	Backbone	1.6E-04	0.009	-1.5	tRNA guanine-N 1-methyltransferase
truB	Z4527	Backbone	1.6E-03	0.030	-1.1	tRNA pseudouridine synthase B
tuf		Backbone	4.7E-03	0.056	-1.4	Elongation factor Tu
ubiB	Z5365	Backbone	3.2E-03	0.043	-1.1	NAD P H-flavin reductase
wbdP	Z3199	O-Island	5.2E-03	0.060	-1.3	Glycosyl transferase
wzy	Z3203	O-Island	1.3E-02	0.096	-1.1	O antigen polymerase
yaeE	Z0210	Backbone	1.3E-03	0.027	-1.4	Putative transport system permease protein
yaeE	Z0211	Backbone	1.2E-02	0.092	-1.3	ATP-binding component of a transporter
yajD	Z0511	Backbone	1.3E-02	0.097	-1.1	Hypothetical protein
ybbA	Z0648	Backbone	5.5E-06	0.003	-1.6	Putative ATP-binding component of a transport system
ybbY	Z0668	Backbone	2.3E-03	0.038	-1.1	Putative transport

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
ybeR	Z0794	Backbone	2.8E-03	0.042	-1.1	Hypothetical protein
yciL	Z2541	Backbone	2.5E-03	0.040	-1.1	Hypothetical protein
yciM	Z2526	Backbone	3.3E-03	0.044	-1.0	Putative heat shock protein
ydcN	Z2285	Backbone	9.1E-03	0.083	-1.0	Hypothetical protein
yehX	Z3378	Backbone	1.2E-02	0.095	-1.1	Putative ATP-binding component of a transport system
yehY	Z3379	Backbone	7.0E-03	0.072	-1.1	Putative transport system permease protein
ygcN	Z4076	Backbone	1.2E-02	0.094	-1.1	Hypothetical protein
yihK	Z5407	Backbone	9.3E-03	0.083	-1.1	Putative GTP-binding factor
yjcZ	Z5712	Backbone	1.0E-02	0.088	-1.1	Hypothetical protein
ylbA	Z0670	Backbone	6.5E-03	0.068	-1.0	Hypothetical protein
Z0330	Z0330	O-Island	1.0E-02	0.086	-1.1	Hypothetical protein
Z0462	Z0462	O-Island	6.5E-03	0.069	-1.0	Putative sensor kinase hexosephosphate transport
Z0609	Z0609	O-Island	3.1E-04	0.012	-1.1	Hypothetical protein
Z0638	Z0638	O-Island	1.2E-02	0.094	-1.0	Hypothetical protein
Z0655	Z0655	Backbone	2.1E-05	0.003	-1.4	Hypothetical protein
Z0891	Z0891	O-Island	1.3E-02	0.096	-1.1	Hypothetical protein
Z1431	Z1431	O-Island	7.1E-03	0.072	-1.1	Hypothetical protein
Z1485	Z1485	O-Island	3.9E-03	0.049	-1.2	Hypothetical protein
Z1500	Z1500	O-Island	9.5E-04	0.022	-1.1	Hypothetical protein
Z1510	Z1510	Backbone	3.3E-03	0.044	-1.1	Putative synthetase
Z1537	Z1537	O-Island	6.4E-03	0.068	-1.1	Putative chaperone
Z1548	Z1548	O-Island	7.4E-03	0.073	-1.1	Putative aminomethyltransferase
Z1784	Z1784	O-Island	1.4E-02	0.099	-1.1	Hypothetical protein
Z3206	Z3206	O-Island	1.3E-02	0.097	-1.1	Putative UDP-galactose 4-epimerase
Z3235m		Backbone	1.2E-02	0.095	-1.1	
Z3359	Z3359	O-Island	6.8E-03	0.070	-1.1	Hypothetical protein

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
Z3625	Z3625	O-Island	9.8E-03	0.085	-1.0	Sucrose hydrolase
Z3635	Z3635	Backbone	3.0E-03	0.043	-1.1	Putative receptor protein
Z3651	Z3651	Backbone	1.1E-02	0.090	-1.0	Putative peptidase
Z3672	Z3672	Junction	7.3E-03	0.072	-1.0	Hypothetical protein
Z4733	Z4733	Backbone	1.4E-02	0.099	-1.1	Hypothetical protein
Z5111	Z5111	O-Island	8.6E-03	0.080	-1.1	Hypothetical protein
Z5881	Z5881	O-Island	5.9E-03	0.065	-1.1	Hypothetical protein
Z5900	Z5900	O-Island	4.5E-03	0.055	-1.1	Hypothetical protein
Z6012	Z6012	O-Island	1.2E-02	0.095	-1.1	Hypothetical protein
Z6072	Z6072	O-Island	4.3E-03	0.053	-1.1	Hypothetical protein
Up-Regulated						
ackA	Z3558	Backbone	1.5E-05	0.003	1.9	Acetate propionate kinase
adhE	Z2016	Backbone	4.3E-04	0.015	1.6	CoA-linked acetaldehyde dehydrogenase and iron-
adiY	Z5718	Backbone	2.4E-04	0.010	1.1	Putative ARAC-type regulatory protein
alaS	Z3999	Backbone	7.2E-04	0.019	1.2	Alanyl-tRNA synthetase
ansB	Z4302	Backbone	7.3E-03	0.072	1.3	Periplasmic L-asparaginase II
aphA	Z5654	Backbone	1.4E-02	0.099	1.1	Diadenosine tetraphosphatase
appA	Z1397	Backbone	1.4E-03	0.028	1.1	Phosphoanhydride phosphorylase pH 2.5 acid phosphatase
appB	Z1396	Backbone	2.3E-05	0.003	1.2	Probable third cytochrome oxidase subunit II
argE	Z5515	Backbone	2.4E-03	0.039	1.2	Acetylornithine deacetylase
aroD	Z2721	Backbone	7.0E-04	0.019	1.1	3-dehydroquinate dehydratase
aroK	Z4743	Backbone	5.8E-04	0.018	1.1	Shikimate kinase I
asd	Z4797	Backbone	1.9E-03	0.033	1.3	Aspartate-semialdehyde dehydrogenase
aspA	Z5744	Backbone	5.6E-05	0.005	2.0	Aspartate ammonia-lyase aspartase
bioF	Z0995	Backbone	2.8E-03	0.042	1.1	8-amino-7-oxononanoate synthase

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
cadA	Z5734	Backbone	3.1E-03	0.043	1.3	Lysine decarboxylase 1
caiA	Z0045	Backbone	9.9E-03	0.085	1.1	Crotonobetainyl-CoA dehydrogenase
caiB	Z0044	Backbone	9.1E-04	0.022	1.0	Crotonobetainyl-CoA:carnitineCoA-transferase
caiE	Z0041	Backbone	9.4E-03	0.084	1.1	Possible synthesis of cofactor for carnitine racemase and
cfa	Z2686	Backbone	5.5E-03	0.063	1.2	Cyclopropane fatty acyl phospholipid synthase
crr	Z3683	Backbone	4.6E-04	0.016	1.3	Glucose-specific PTS system enzyme IIA component
csgA	Z1676	Backbone	5.4E-04	0.018	1.9	Curlin major subunit coiled surface structures cryptic
csgB	Z1675	Backbone	1.8E-04	0.009	1.4	Minor curlin subunit precursor similar to CsgA
csgC	Z1677	Backbone	1.2E-04	0.008	1.5	Putative curli production protein
csgD	Z1673	Backbone	1.0E-02	0.086	1.2	Putative 2-component transcriptional regulator for 2nd curli operon
csgE	Z1672	Backbone	2.5E-03	0.040	1.3	Curli assembly transport component 2nd curli operon
csgF	Z1671	Backbone	1.3E-02	0.096	1.5	Curli assembly transport component 2nd curli operon
csgG	Z1670	Backbone	4.0E-03	0.051	1.4	Curli assembly transport component 2nd curli operon
cysE	Z5034	Backbone	1.5E-04	0.008	1.1	Serine acetyltransferase
cysK	Z3680	Backbone	1.4E-03	0.027	1.3	Cysteine synthase A O-acetylserine sulfhydrolase A
dacC	Z1066	Backbone	3.1E-03	0.043	1.0	D-alanyl-D-alanine carboxypeptidase penicillin-binding
dcuB	Z5725	Backbone	3.1E-03	0.043	1.0	Anaerobic C4-dicarboxylate transporter
dnaE	Z0196	Backbone	6.8E-03	0.070	1.1	DNA polymerase III subunit alpha
dnaJ	Z0015	Backbone	8.3E-04	0.020	1.2	Chaperone with DnaK heat shock protein
dnaK	Z0014	Backbone	3.1E-04	0.012	2.0	Molecular chaperone DnaK
dnaX	Z0587	Backbone	7.1E-03	0.072	1.1	DNA polymerase III subunits gamma and tau
dps	Z1034	Backbone	1.1E-02	0.089	1.2	DNA protection during starvation conditions
dsbD	Z5741	Backbone	8.3E-03	0.079	1.1	Thiol:disulfide interchange protein precursor
dxs	Z0523	Backbone	8.9E-03	0.082	1.1	1-deoxy-D-xylulose-5-phosphate synthase
eaeH	Z0375	Junction	5.5E-03	0.063	1.0	Putative adhesin
elaB	Z3526	Backbone	1.2E-03	0.027	1.2	Hypothetical protein

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
emrD	Z5168	Backbone	4.4E-03	0.055	1.1	2-module integral membrane pump multidrug resistance
eno	Z4094	Backbone	7.8E-04	0.020	1.5	Phosphopyruvate hydratase
escR	Z5135	O-Island	8.4E-03	0.079	1.1	Type III secretion system protein
evgA	Z3631	Backbone	1.6E-03	0.030	1.1	Putative positive transcription regulator sensor EvgS
fba	Z4263	Backbone	1.7E-03	0.031	1.4	Fructose-bisphosphate aldolase
fdhF	Z5678	Backbone	1.5E-05	0.003	1.4	Formate dehydrogenase
fhlA	Z4040	Backbone	2.9E-03	0.042	1.2	Formate hydrogen-lyase transcriptional activator for fdhF
flgE	Z1714	Junction	7.9E-03	0.076	1.1	Flagellar hook protein
fmt	Z4658	Backbone	4.0E-03	0.051	1.0	Methionyl-tRNA formyltransferase
ftsJ	Z4541	Backbone	2.2E-04	0.010	1.3	Cell division protein
gadA	Z4930	Backbone	1.1E-04	0.007	3.6	Glutamate decarboxylase isozyme
galK	Z0927	Backbone	1.1E-02	0.090	1.1	Galactokinase
galM	Z0926	Backbone	1.8E-04	0.009	1.3	Galactose-1-epimerase mutarotase
galP	Z4288	Backbone	2.2E-03	0.038	1.1	Galactose-proton symport of transport system
gapA	Z2818	Backbone	3.1E-03	0.043	1.1	Glyceraldehyde-3-phosphate dehydrogenase
gapC	Z2304	Backbone	2.0E-04	0.009	1.1	Glyceraldehyde-3-phosphate dehydrogenase
gcvH	Z4241	Backbone	7.3E-03	0.072	1.4	Glycine cleavage system protein H
gcvP	Z4240	Backbone	3.2E-03	0.043	1.3	Glycine dehydrogenase
gcvR	Z3738	Backbone	9.7E-03	0.085	1.1	Transcriptional regulation of gcv operon
gdhA	Z2793	Backbone	2.5E-03	0.040	1.1	Glutamate dehydrogenase
glgA	Z4791	Backbone	3.8E-03	0.049	1.1	Glycogen synthase
glgB	Z4796	Backbone	4.6E-03	0.056	1.1	Glycogen branching enzyme
glgC	Z4792	Backbone	7.3E-03	0.072	1.2	Glucose-1-phosphate adenylyltransferase
glgP	Z4790	Junction	4.3E-04	0.015	1.3	Glycogen phosphorylase
glgX	Z4794	Backbone	3.1E-03	0.043	1.1	Glycogen debranching enzyme
glnE	Z4406	Backbone	4.5E-03	0.055	1.0	Adenylylating enzyme for glutamine synthetase

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
gpsA	Z5035	Backbone	3.9E-05	0.004	1.2	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase
groEL	Z5748	Backbone	1.1E-04	0.007	1.5	Chaperonin GroEL
groES	Z5747	Backbone	1.1E-04	0.007	1.7	Co-chaperonin GroES
grxB	Z1701	Backbone	6.0E-04	0.019	1.4	Glutaredoxin 2
gshB	Z4292	Backbone	1.1E-03	0.025	1.2	Glutathione synthetase
gsk	Z0596	Backbone	6.3E-04	0.019	1.1	Inosine-guanosine kinase
guaC	Z0114	Backbone	1.2E-03	0.027	1.2	Guanosine 5%27-monophosphate oxidoreductase
hdeA	Z4922	Backbone	3.3E-05	0.004	2.3	Hypothetical protein
hflB	Z4540	Backbone	7.1E-04	0.019	1.2	Degrades sigma32
hflC	Z5782	Backbone	9.1E-03	0.083	1.2	Protease specific for phage lambda cII repressor
hflK	Z5781	Backbone	4.1E-04	0.015	1.2	Protease specific for phage lambda cII repressor
hflX	Z5780	Backbone	2.4E-03	0.039	1.3	GTP - binding subunit of protease specific for phage
hfq	Z5779	Backbone	4.9E-04	0.016	1.7	RNA-binding protein Hfq
hisS	Z3777	Backbone	5.4E-03	0.061	1.2	Histidyl-tRNA synthetase
hslU	Z5478	Backbone	7.7E-05	0.006	1.3	ATP-dependent protease ATP-binding subunit
hslV	Z5479	Backbone	1.8E-04	0.009	1.5	ATP-dependent protease peptidase subunit
hyaA	Z1389	Backbone	9.2E-04	0.022	1.1	Hydrogenase-1 small subunit
hyaB	Z1390	Backbone	7.1E-03	0.072	1.2	Hydrogenase-1 large subunit
hyaC	Z1391	Backbone	1.5E-03	0.029	1.1	Probable Ni Fe-hydrogenase 1 b-type cytochrome subunit
hyaF	Z1394	Backbone	2.9E-03	0.043	1.1	Nickel incorporation into hydrogenase-1 proteins
hycA	Z4033	Backbone	3.6E-04	0.014	1.2	Transcriptional repression of hyc and hyp operons
hycB	Z4032	Backbone	2.8E-03	0.042	1.1	Probable small subunit of hydrogenase-3 iron-sulfur protein
hycD	Z4030	Backbone	9.2E-03	0.083	1.2	Membrane-spanning protein of hydrogenase 3 part of FHL
hycE	Z4029	Backbone	9.9E-04	0.023	1.2	Large subunit of hydrogenase 3 part of FHL complex
hycF	Z4028	Backbone	9.6E-05	0.007	1.5	Hydrogenase 4 Fe-S subunit
hycG	Z4027	Backbone	5.4E-04	0.018	1.8	Hydrogenase activity

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
hycH	Z4026	Backbone	5.6E-04	0.018	1.2	Processing of large subunit HycE of hydrogenase 3 part of
hycI	Z4025	Backbone	2.0E-04	0.009	1.1	Hydrogenase 3 maturation protease
hypA	Z4035	Backbone	6.3E-04	0.019	1.1	Hydrogenase nickel incorporation protein
hypA	Z4345	Backbone	6.3E-04	0.019	1.1	Hydrogenase nickel incorporation protein
hypB	Z4036	Backbone	6.2E-04	0.019	1.2	Guanine-nucleotide binding protein functions as nickel
hypE	Z4039	Backbone	2.8E-03	0.042	1.2	Plays structural role in maturation of all 3 hydrogenase
icc	Z4389	Backbone	2.6E-03	0.041	1.1	Regulator of lacZ
ileS	Z0030	Backbone	1.3E-02	0.099	1.1	Isoleucyl-tRNA synthetase
intT	Z2966	O-Island	1.2E-02	0.094	1.1	Integrase for prophage CP-933T
kefC	Z0053	Backbone	1.1E-02	0.092	1.1	Glutathione-regulated potassium-efflux system protein
lpxA	Z0193	Backbone	1.3E-02	0.096	1.1	UDP-N-acetylglucosamine acyltransferase
lysU	Z5732	Backbone	7.6E-06	0.003	1.9	Lysyl-tRNA synthetase
manY	Z2861	Backbone	5.8E-03	0.064	1.1	PTS enzyme IIC mannose-specific
manZ	Z2862	Backbone	2.4E-03	0.039	1.1	PTS enzyme IID mannose-specific
mdaB	Z4379	Backbone	3.3E-03	0.044	1.3	Modulator of drug activity B
mdoG	Z1683	Backbone	5.8E-03	0.064	1.3	Periplasmic glucans biosynthesis protein
miaA	Z5778	Backbone	6.4E-03	0.068	1.5	tRNA delta 2 -isopentenylpyrophosphate transferase
mipB	Z1048	Backbone	1.4E-03	0.027	1.1	Fructose-6-phosphate aldolase
murB	Z5543	Backbone	1.1E-04	0.007	1.3	UDP-N-acetylenolpyruvoylglucosamine reductase
murI	Z5528	Backbone	2.3E-03	0.039	1.2	Glutamate racemase
narZ	Z2244	Backbone	1.3E-02	0.097	1.1	Cryptic nitrate reductase 2 alpha subunit
nhaR	Z0019	Backbone	1.5E-03	0.029	1.1	Transcriptional activator of nhaA
nlpD	Z4050	Backbone	1.4E-05	0.003	2.1	Lipoprotein
nrfB	Z5670	Backbone	8.1E-03	0.077	1.2	Formate-dependent nitrite reductase
nrfC	Z5671	Backbone	5.8E-03	0.064	1.2	Formate-dependent nitrite reductase
nrfD	Z5672	Backbone	5.8E-06	0.003	1.2	Formate-dependent nitrate reductase complex transmembrane protein

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
ompA	Z1307	Backbone	8.7E-04	0.021	1.4	Outer membrane protein 3a II G d
oppA	Z2019	Backbone	2.2E-04	0.010	1.2	Oligopeptide transport periplasmic binding protein
oppB	Z2020	Backbone	1.4E-02	0.099	1.1	Oligopeptide permease ABC transporter membrane component
oppD	Z2022	Backbone	1.7E-03	0.031	1.1	Oligopeptide transporter ATP-binding component
otsA	Z2949	Backbone	1.0E-03	0.023	1.2	Trehalose-6-phosphate synthase
pcnB	Z0154	Backbone	1.3E-02	0.096	1.2	Poly A polymerase I
pdxK	Z3684	Backbone	1.0E-02	0.087	1.4	Pyridoxine kinase
pepD	Z0298	Backbone	2.8E-05	0.003	1.3	Aminoacyl-histidine dipeptidase peptidase D
pepE	Z5612	Backbone	5.8E-03	0.064	1.2	Peptidase E
pfkB	Z2752	Backbone	1.3E-03	0.027	1.1	6-phosphofructokinase II suppressor of pfkA
pgi	Z5623	Backbone	6.1E-05	0.005	1.2	Glucose-6-phosphate isomerase
pgk	Z4265	Backbone	1.0E-02	0.086	1.3	Phosphoglycerate kinase
pphB	Z4044	Junction	3.7E-04	0.014	1.2	Protein phosphatase 2
pqiA	Z1300	Backbone	1.3E-02	0.097	1.0	Paraquat-inducible protein A
pta	Z3559	Backbone	1.0E-05	0.003	2.1	Phosphate acetyltransferase
pth	Z1975	Backbone	6.0E-03	0.065	1.1	Peptidyl-tRNA hydrolase
ptsI	Z3682	Backbone	2.6E-04	0.011	1.2	Phosphoenolpyruvate-protein phosphotransferase
qseB	Z4377	Backbone	2.1E-04	0.010	1.1	DNA-binding transcriptional regulator QseB
ribB	Z4399	Backbone	3.4E-04	0.013	1.6	3 4-dihydroxy-2-butanone 4-phosphate synthase
ribF	Z0029	Backbone	1.2E-02	0.092	1.1	Hypothetical protein
rpoN	Z4565	Backbone	2.2E-03	0.038	1.1	DNA-directed RNA polymerase subunit N
rpoS	Z4049	Backbone	2.3E-05	0.003	2.0	RNA polymerase sigma factor
rpsB	Z0180	Backbone	6.2E-03	0.067	1.1	30S ribosomal protein S2
secG	Z4537	Backbone	9.1E-03	0.083	1.1	Protein-export membrane protein
speD	Z0130	Backbone	2.0E-03	0.036	1.2	S-adenosylmethionine decarboxylase proenzyme

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
speF	Z0839	Backbone	2.4E-03	0.039	1.2	Ornithine decarboxylase isozyme inducible
spr	Z3434	Backbone	1.8E-03	0.032	1.5	Putative lipoprotein
talA	Z3720	Backbone	1.1E-04	0.007	1.5	Transaldolase
tdcA	Z4470	Backbone	5.7E-03	0.064	1.2	Transcriptional activator of tdc operon
tdcB	Z4469	Backbone	1.1E-04	0.007	2.3	Threonine dehydratase
tdcC	Z4468	Backbone	3.2E-05	0.004	1.6	Anaerobically inducible L-threonine L-serine permease
tdcD	Z4467	Backbone	6.6E-03	0.069	1.1	Acetate propionate kinase
tdcE	Z4466	Backbone	3.2E-04	0.012	2.7	Probable formate acetyltransferase 3
tdk	Z2015	Backbone	5.2E-04	0.017	1.3	Thymidine kinase
thrA	Z0002	Backbone	3.3E-03	0.044	1.1	Bifunctional aspartokinase I homeserine dehydrogenase I
thrC	Z0004	Backbone	2.9E-04	0.012	1.0	Threonine synthase
tktA	Z4279	Backbone	6.8E-04	0.019	1.3	Transketolase
tolC	Z4392	Backbone	7.8E-03	0.075	1.2	Outer membrane channel precursor protein
tpx	Z2452	Backbone	3.8E-03	0.049	1.1	Thiol peroxidase
tufB	Z5553	Backbone	6.1E-04	0.019	1.2	Elongation factor Tu
ubiE	Z5355	Backbone	7.8E-03	0.075	1.2	Ubiquinone menaquinone biosynthesis methyltransferase
ushA	Z0599	Backbone	8.6E-04	0.021	1.1	UDP-sugar hydrolase 5 -nucleotidase
waaQ	Z5056	Junction	1.2E-02	0.095	1.1	Putative LPS biosynthesis enzyme
yaaF	Z0035	Backbone	3.1E-03	0.043	1.1	Hypothetical protein
yacE	Z0113	Backbone	3.3E-03	0.044	1.2	Putative DNA repair protein
yacG	Z0111	Backbone	5.3E-03	0.061	1.3	Zinc-binding protein
yadB	Z0155	Backbone	1.3E-02	0.096	1.1	Glutamyl-tRNA synthetase
yadM	Z0149	Hypervari	1.1E-02	0.092	1.1	Putative fimbrial protein
yahE	Z0408	Backbone	9.4E-03	0.084	1.1	Hypothetical protein
yahG	Z0410	Backbone	8.7E-03	0.081	1.1	Hypothetical protein
ybaP	Z0601	Backbone	8.1E-03	0.077	1.1	Putative ligase

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
ybaS	Z0606	Backbone	1.3E-03	0.027	1.2	Glutaminase
ybaT	Z0607	Backbone	4.6E-04	0.016	1.2	Putative amino acid amine transport protein
ybbJ	Z0641	Junction	1.1E-02	0.091	1.1	Hypothetical protein
ybgL	Z0864	Backbone	1.2E-02	0.094	1.1	Hypothetical protein
ybhE	Z0938	Backbone	1.3E-03	0.027	1.2	Putative isomerase
ybhF	Z1014	Backbone	2.0E-03	0.034	1.2	Putative ATP-binding component of a transport system
ybhS	Z1013	Backbone	4.7E-03	0.056	1.1	Hypothetical protein
ybiS	Z1041	Backbone	5.9E-03	0.065	1.1	Hypothetical protein
ycbG	Z1306	Backbone	1.2E-03	0.027	1.0	Hypothetical protein
yceK	Z1685	Backbone	1.3E-02	0.096	1.1	Hypothetical protein
ycfF	Z1742	Backbone	7.2E-04	0.019	1.1	Hypothetical protein
ycfM	Z1744	Backbone	7.0E-04	0.019	1.2	Hypothetical protein
ycfN	Z1745	Backbone	6.2E-03	0.067	1.1	Putative beta-glucosidase
ycfP	Z1747	Backbone	1.4E-02	0.099	1.1	Hypothetical protein
yciB	Z2032	Backbone	9.3E-03	0.083	1.1	Putative intracellular septation protein
ydaA	Z2435	Backbone	2.3E-03	0.038	1.2	Hypothetical protein
ydiJ	Z2715	Backbone	1.5E-04	0.008	1.1	Putative oxidase
ydjJ	Z2812	Backbone	4.2E-03	0.052	1.1	Putative oxidoreductase
yeaD	Z2820	Backbone	9.9E-03	0.085	1.1	Hypothetical protein
yeaG	Z2823	Backbone	4.2E-03	0.053	1.1	Hypothetical protein
yeaH	Z2824	Backbone	1.3E-02	0.097	1.2	Hypothetical protein
yeaU	Z2843	Backbone	8.3E-04	0.020	1.2	Putative tartrate dehydrogenase
yecO	Z2923	Backbone	6.3E-03	0.067	1.1	Hypothetical protein
yeeA	Z3169	Backbone	3.6E-03	0.046	1.1	Hypothetical protein
yegN	Z3244	Backbone	5.7E-03	0.064	1.2	Hypothetical protein
yfbB	Z3523	Backbone	6.8E-03	0.071	1.1	Putative enzyme

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
yfbT	Z3554	Backbone	9.3E-05	0.007	1.3	Putative phosphatase
ygaG	Z3988	Backbone	7.8E-03	0.075	1.0	S-ribosylhomocysteinase
ygaM	Z3972	Backbone	1.3E-03	0.027	1.3	Hypothetical protein
ygcP	Z4078	Backbone	8.8E-03	0.082	1.0	Putative anti-terminator regulatory protein
yggE	Z4259	Backbone	1.5E-04	0.008	1.3	Putative actin
ygiW	Z4376	Backbone	3.7E-05	0.004	1.3	Hypothetical protein
yhaH	Z4457	Backbone	1.3E-02	0.097	1.1	Putative cytochrome
yhaR	Z4465	Backbone	1.5E-04	0.008	2.6	Hypothetical protein
yhbG	Z4564	Backbone	7.1E-03	0.072	1.1	Putative ATP-binding component of a transport system
yhbL	Z4573	Backbone	4.5E-03	0.055	1.1	Sigma cross-reacting protein 27A SCRP-27A
yhbN	Z4563	Backbone	2.2E-03	0.037	1.1	Hypothetical protein
yheL	Z4701	Backbone	4.3E-03	0.053	1.1	Hypothetical protein
yheO	Z4704	Backbone	1.3E-04	0.008	1.2	Hypothetical protein
yhfA	Z4717	Backbone	2.6E-03	0.041	1.0	Hypothetical protein
yhgH	Z4768	Backbone	4.9E-03	0.058	1.1	Hypothetical protein
yhhJ	Z4884	Backbone	9.7E-03	0.085	1.1	Putative transporter
yhhX	Z4808	Backbone	1.2E-02	0.093	1.1	Putative regulator
yhiE	Z4925	Backbone	4.9E-03	0.057	1.2	Hypothetical protein
yhiM	Z4890	Backbone	1.5E-03	0.029	1.2	Hypothetical protein
yhiP	Z4896	Backbone	2.7E-03	0.042	1.2	Putative transport protein
yhiU	Z4926	Backbone	3.0E-03	0.043	1.2	Hypothetical protein
yibP	Z5040	Backbone	1.3E-03	0.027	1.2	Hypothetical protein
yjaH	Z5577	Backbone	3.1E-03	0.043	1.1	Hypothetical protein
yjeE	Z5775	Backbone	9.3E-03	0.083	1.1	Hypothetical protein
yjeK	Z5751	Backbone	3.0E-03	0.043	1.1	Hypothetical protein
yjeT	Z5783	Backbone	8.0E-04	0.020	1.1	Hypothetical protein

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
ykgC	Z0381	Backbone	7.1E-04	0.019	1.1	Pyridine nucleotide-disulfide oxidoreductase
ykgI	Z0363	Backbone	2.5E-03	0.040	1.1	Hypothetical protein
ymjA	Z2493	Backbone	1.1E-02	0.089	1.1	Hypothetical protein
yqjC	Z4451	Backbone	1.7E-03	0.031	1.4	Hypothetical protein
yqjD	Z4452	Backbone	6.6E-04	0.019	1.1	Hypothetical protein
yqjE	Z4453	Backbone	3.5E-03	0.046	1.2	Hypothetical protein
yrbK	Z4562	Backbone	7.7E-03	0.075	1.1	Hypothetical protein
yrfG	Z4753	Backbone	7.1E-03	0.072	1.1	Putative phosphatase
Z0202	Z0202	Backbone	2.8E-03	0.042	1.1	Hypothetical protein
Z0370	Z0370	O-Island	2.2E-03	0.038	1.1	Hypothetical protein
Z0373	Z0373	O-Island	4.8E-04	0.016	1.0	Hypothetical protein
Z0379	Z0379	Backbone	6.8E-04	0.019	1.1	Hypothetical protein
Z0402	Z0402	O-Island	5.0E-03	0.058	1.1	Putative beta-barrel outer membrane protein
Z0615	Z0615	O-Island	5.5E-04	0.018	1.2	Putative RTX family exoprotein
Z0634	Z0634	O-Island	2.0E-05	0.003	1.5	Putative cytoplasmic membrane export protein
Z0952	Z0952	O-Island	1.3E-03	0.027	1.1	Putative Bet recombination protein of prophage CP-933K
Z0968	Z0968	O-Island	1.4E-04	0.008	2.2	Hypothetical protein
Z1015	Z1015	Backbone	3.2E-03	0.043	1.1	Hypothetical protein
Z1020	Z1020	Junction	9.7E-03	0.085	1.1	Putative ATP-dependent helicase
Z1336	Z1336	O-Island	8.9E-03	0.082	1.1	Hypothetical protein
Z1442	Z1442	O-Island	1.1E-02	0.089	1.1	Putative antitermination protein N of bacteriophage BP-
Z1444	Z1444	O-Island	5.9E-03	0.065	1.5	Putative serine threonine kinase encoded by bacteriophage
Z1453	Z1453	O-Island	8.9E-03	0.082	1.4	Hypothetical protein
Z1498	Z1498	O-Island	4.7E-03	0.056	1.6	Hypothetical protein
Z1648	Z1648	O-Island	6.5E-03	0.068	1.1	Unknown in putative ISEc8
Z1679	Z1679	Backbone	4.1E-03	0.051	1.1	Hypothetical protein

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
Z1890	Z1890	O-Island	9.8E-03	0.085	1.1	Putative head-tail joining protein of prophage CP-933X
Z2206	Z2206	Backbone	1.3E-02	0.096	1.1	Putative adhesin similar to FimH protein
Z2221	Z2221	Backbone	1.1E-02	0.090	1.4	Hypothetical protein
Z2272	Z2272	Backbone	1.1E-02	0.089	1.0	Hypothetical protein
Z2275	Z2275	Backbone	1.2E-02	0.095	1.1	Putative aldehyde dehydrogenase
Z2319	Z2319	Backbone	2.1E-03	0.036	1.3	Probable enzyme
Z2323	Z2323	O-Island	5.8E-03	0.064	1.1	Hypothetical protein
Z2393	Z2393	Backbone	1.1E-02	0.092	1.1	Hypothetical protein
Z2398	Z2398	Backbone	5.1E-03	0.059	1.1	Hypothetical protein
Z2399	Z2399	Backbone	2.7E-03	0.042	1.1	Putative regulatory protein Cro of prophage CP-933R
Z2400	Z2400	Backbone	2.9E-04	0.012	1.6	Partial putative repressor protein encoded within prophage CP-933R
Z2469	Z2469	Backbone	5.0E-03	0.058	1.0	Putative transient receptor potential locus
Z2581	Z2581	Backbone	7.3E-03	0.072	1.1	Putative oxidoreductase component
Z2753	Z2753	Backbone	2.6E-03	0.041	1.2	Hypothetical protein
Z2879	Z2879	Backbone	7.6E-03	0.074	1.1	Hypothetical protein
Z2968	Z2968	O-Island	1.2E-02	0.092	1.1	Hypothetical protein
Z3087	Z3087	O-Island	1.1E-02	0.090	1.1	Putative tail fiber component V of prophage CP-933U
Z3116	Z3116	O-Island	7.0E-03	0.072	1.0	Hypothetical protein
Z3332	Z3332	O-Island	8.5E-04	0.021	1.5	Hypothetical protein
Z3395	Z3395	O-Island	3.2E-03	0.043	1.0	Putative regulator
Z3551	Z3551	Backbone	1.5E-03	0.029	1.1	Aspartate aminotransferase
Z3561	Z3561	Backbone	9.4E-04	0.022	1.2	Putative regulator
Z3919	Z3919	O-Island	1.2E-02	0.094	1.1	Hypothetical protein
Z4165	Z4165	Backbone	9.2E-03	0.083	1.2	Putative transporter protein
Z4464	Z4464	Backbone	3.5E-05	0.004	1.8	Putative L-serine deaminase
Z4861	Z4861	O-Island	3.5E-03	0.046	1.1	Hypothetical protein

Table 3.S1. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
Z5138	Z5138	O-Island	1.3E-02	0.096	1.1	Hypothetical protein
Z5140	Z5140	O-Island	9.7E-03	0.085	1.1	Hypothetical protein
Z5547	Z5547	Backbone	2.0E-03	0.034	1.2	Hypothetical protein
Z5766	Z5766	Backbone	4.6E-03	0.056	1.1	Phosphatidylserine decarboxylase
Z5898	Z5898	O-Island	1.0E-02	0.086	1.2	Hypothetical protein
Z5947	Z5947	Hyper-	1.4E-02	0.099	1.0	Putative restriction modification enzyme M subunit
Z6010	Z6010	O-Island	1.0E-02	0.087	1.3	Hypothetical protein

APPENDIX C

TABLE 3.S2. DIFFERENTIALLY EXPRESSED GENES OF *E. COLI* O157:H7 DURING RUMEN FLUID TREATMENT

Table 3.S2. Differentially expressed genes of *E. coli* O157:H7 during rumen fluid treatment

Gene	Locus	Location	*p* value	q value	FC	Function/Product
Down-regulated						
agaI_1	Z4495	Backbone	5.0E-04	0.049	-1.1	
argS	Z2929	Backbone	1.6E-03	0.090	-1.3	Arginyl-tRNA synthetase
fabF	Z1734	Backbone	1.6E-05	0.018	-1.6	3-oxoacyl-(acyl carrier protein) synthase II
frdA	Z5762	Backbone	7.5E-04	0.061	-1.6	Fumarate reductase flavoprotein subunit
fusA	Z4698	Backbone	3.1E-04	0.036	-2.7	Elongation factor G
gpt	Z0299	Backbone	1.3E-03	0.085	-1.4	Xanthine-guanine phosphoribosyltransferase
hemK	Z1983	Backbone	1.4E-03	0.085	-1.1	N5-glutamine S-adenosyl-L-methionine-dependent
hisS	Z3777	Backbone	1.3E-03	0.085	-1.4	Histidyl-tRNA synthetase
nadC	Z0119	Backbone	1.3E-03	0.085	-1.5	Quinolinate phosphoribosyltransferase
ompA	Z1307	Backbone	4.0E-04	0.044	-1.6	Outer membrane protein A
ompX	Z1036	Backbone	1.4E-03	0.085	-1.3	Outer membrane protein X
prsA	Z1978	Backbone	3.6E-06	0.011	-2.1	Ribose-phosphate pyrophosphokinase
pta	Z3559	Backbone	1.8E-05	0.018	-4.1	Phosphate acetyltransferase
ptsG	Z1740	Backbone	3.0E-04	0.036	-2.3	Glucose-specific PTS system IIBC components
rplC	Z4691	Backbone	5.0E-04	0.049	-2.8	50S ribosomal protein L3
rplD	Z4690	Backbone	1.4E-03	0.087	-2.7	50S ribosomal protein L4
rplP	Z4684	Backbone	1.4E-03	0.085	-1.9	50S ribosomal protein L16
rplV	Z4686	Backbone	7.0E-05	0.024	-2.6	50S ribosomal protein L22
rpsC	Z4685	Backbone	1.8E-04	0.035	-1.8	30S ribosomal protein S3
rpsG	Z4699	Backbone	8.4E-04	0.066	-3.6	30S ribosomal protein S7
rpsH	Z4676	Backbone	2.0E-03	0.100	-1.9	30S ribosomal protein S8
rpsK	Z4667	Backbone	7.5E-05	0.024	-2.8	30S ribosomal protein S11
rpsS	Z4687	Backbone	1.2E-03	0.085	-2.4	30S ribosomal protein S19
tdH	Z5043	Backbone	1.3E-03	0.085	-1.3	L-threonine 3-dehydrogenase
tehB	Z2288	Backbone	1.3E-03	0.085	-1.1	Tellurite resistance protein TehB
thyA	Z4144	Backbone	7.7E-04	0.061	-1.3	Thymidylate synthase

Table 3.S2. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
tufA	Z4697	Backbone	2.7E-04	0.035	-2.9	Elongation factor Tu
tyrP	Z2963	Backbone	1.6E-03	0.090	-1.1	Tyrosine-specific transport system
uhpA	Z5159	Backbone	3.1E-04	0.036	-1.2	DNA-binding transcriptional activator UhpA
uup	Z1299	Backbone	5.3E-04	0.049	-1.3	ABC transporter ATPase component
waaL	Z5049	O-island	4.3E-04	0.046	-1.3	Putative LPS biosynthesis enzyme
yaaH	Z0010	Backbone	1.4E-03	0.087	-1.2	Hypothetical protein
ybbA	Z0648	Backbone	3.3E-05	0.021	-2.2	Putative ABC transporter ATP-binding protein YbbA
yfhJ	Z3791	Backbone	6.4E-04	0.057	-1.2	Hypothetical protein
yhbE	Z4546	Backbone	1.9E-03	0.100	-2.2	Hypothetical protein
yibN	Z5038	Backbone	1.2E-04	0.031	-1.5	Hypothetical protein
yibO	Z5039	Backbone	1.7E-04	0.035	-2.1	Phosphoglyceromutase
yiiR	Z5466	Backbone	6.8E-04	0.059	-1.1	Hypothetical protein
yojN	Z3475	Backbone	4.7E-05	0.021	-1.8	Phosphotransfer intermediate proteinwith RcsBC
Z0110	Z0110	Backbone	2.2E-05	0.019	-1.2	Hypothetical protein
Z1506	Z1506	Backbone	1.8E-04	0.035	-1.3	Putative 4-hydroxyphenylacetate 3-monooxygenase
Z1878	Z1878	O-island	1.3E-03	0.085	-1.3	Putative Bor protein of prophage CP-933X
Z4351	Z4351	Backbone	2.7E-04	0.035	-2.0	Hydrogenase 2 small subunit
Z4875	Z4875	O-island	1.7E-03	0.092	-1.8	Putative phosphotransferase system enzyme subunit
Z5140	Z5140	O-island	1.6E-03	0.090	-1.2	Hypothetical protein
Up-regulated						
adiA	Z5719	Backbone	5.4E-06	0.011	2.4	Biodegradative arginine decarboxylase
aspA	Z5744	Backbone	2.0E-04	0.035	3.8	Aspartate ammonia-lyase
cfa	Z2686	Backbone	1.7E-04	0.035	1.9	Cyclopropane fatty acyl phospholipid synthase
csgA	Z1676	Backbone	2.0E-03	0.100	1.5	Cryptic curlin major subunit
cydB	Z0901	Backbone	9.4E-05	0.028	1.3	Cytochrome d terminal oxidase polypeptide subunit II
cysK	Z3680	Backbone	5.6E-05	0.021	4.6	Cysteine synthase A
dnaK	Z0014	Backbone	5.4E-04	0.049	2.0	Molecular chaperone DnaK

Table 3.S2. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
dps	Z1034	Backbone	2.7E-04	0.035	1.6	DNA starvation/stationary phase protection protein Dps
focA	Z1250	Backbone	1.6E-03	0.090	1.2	Formate transporter
gadA	Z4930	Backbone	4.9E-05	0.021	3.2	Glutamate decarboxylase isozyme
galM	Z0926	Backbone	1.7E-03	0.092	1.2	Aldose 1-epimerase
hfq	Z5779	Backbone	2.2E-04	0.035	1.3	RNA-binding protein Hfq
hyaA	Z1389	Backbone	5.2E-04	0.049	1.3	Hydrogenase-1 small subunit
lysU	Z5732	Backbone	2.4E-04	0.035	1.5	Lysyl-tRNA synthetase
malF	Z5631	Backbone	2.7E-04	0.035	6.0	Maltose transporter membrane protein
narG	Z2001	Backbone	1.5E-03	0.087	1.2	Nitrate reductase 1, alpha subunit
nlpD	Z4050	Backbone	2.2E-04	0.035	3.1	Lipoprotein NlpD
pckA	Z4758	Backbone	2.0E-03	0.100	1.6	Phosphoenolpyruvate carboxykinase
rimJ	Z1703	Backbone	1.0E-03	0.076	1.1	Ribosomal-protein-S5-alanine N-acetyltransferase
rpoS	Z4049	Backbone	2.2E-04	0.035	2.0	RNA polymerase sigma factor RpoS
talB	Z0008	Backbone	3.2E-04	0.036	1.5	Transaldolase B
uspE	Z2435	Backbone	5.0E-04	0.049	3.3	Universal stress protein UspE
vacB	Z5786	Backbone	2.6E-04	0.035	1.7	Exoribonuclease R, RNase R
yabI	Z0074	Backbone	1.3E-03	0.085	1.1	Hypothetical protein
ycbO	Z1284	Backbone	2.0E-03	0.100	1.2	Alkanesulfonate transporter substrate-binding subunit
yceP	Z1697	Backbone	5.1E-05	0.021	1.7	Biofilm formation regulatory protein BssS
ydfG	Z2158	Backbone	2.0E-03	0.100	1.1	3-hydroxy acid dehydrogenase
yeaU	Z2843	Backbone	2.7E-04	0.035	4.0	Putative tartrate dehydrogenase
ygaM	Z3972	Backbone	5.2E-04	0.049	1.3	Hypothetical protein
ygdH	Z4112	Backbone	1.1E-04	0.030	1.4	Hypothetical protein
yhaH	Z4457	Backbone	1.8E-04	0.035	1.3	Putative cytochrome
yhaR	Z4465	Backbone	1.8E-04	0.035	3.3	Hypothetical protein
yojL	Z3472	Backbone	7.5E-04	0.061	1.1	Thiamine biosynthesis lipoprotein ApbE
Z1490	Z1490	O-island	7.0E-04	0.059	1.1	Hypothetical protein

Table 3.S2. (continued)

Gene	Locus	Location	*p* value	q value	FC	Function/Product
Z1491	Z1491	O-island	9.5E-04	0.073	1.1	Hypothetical protein
Z2353	Z2353	Backbone	3.4E-05	0.021	2.9	Putative tail component of prophage CP-933R
Z2398	Z2398	Backbone	1.3E-03	0.085	1.2	Hypothetical protein
Z5001	Z5001	O-island	1.6E-03	0.090	1.2	Putative permease

APPENDIX D

TABLE 4.S1. DIFFERENTIALLY EXPRESSED GENES OF *E. COLI* O157:H7 WITHIN *ACANTHAMOEBA*

Table 4.S1. Differentially expressed genes of *E. coli* O157:H7 within *Acanthamoeba*

ID	Locus	Function/Product	*p value*	q value	FC
Down-Regulated					
accB	Z4615	Acetyl CoA carboxylase, BCCP subunit	3.94E-02	2.49E-02	-1.4
acnA	Z2532	Aconitate hydratase 1	8.51E-03	9.21E-03	-1.4
adhE	Z2016	Multifunctional acetaldehyde-CoA dehydrogenase and iron-dependent alcohol dehydrogenase and pyruvate-formate lyase deactivase	6.79E-06	2.17E-04	-2.3
adiA	Z5719	Biodegradative arginine decarboxylase	9.72E-05	9.59E-04	-3.1
adiY	Z5718	DNA-binding transcriptional activator	2.38E-03	4.82E-03	-1.5
appA	Z1397	Phosphoanhydride phosphorylase	1.59E-04	1.25E-03	-1.3
appB	Z1396	Cytochrome bd-II oxidase, subunit II	3.29E-05	4.96E-04	-1.8
araH_B	Z2951	Partial high-affinity L-arabinose transport system; membrane protein, fragment 2	7.21E-04	2.71E-03	-1.4
asd	Z4797	Aspartate-semialdehyde dehydrogenase, NAD(P)-binding	2.46E-05	4.37E-04	-1.7
aspA	Z5744	Aspartate ammonia-lyase	8.22E-09	2.11E-06	-2.1
baeR	Z3248	DNA-binding response regulator in two-component regulatory system with BaeS	5.06E-03	7.03E-03	-1.3
bcsE	Z4952	Conserved protein	1.12E-04	1.04E-03	-1.4
bcsF	Z4953	Predicted protein	1.94E-02	1.53E-02	-1.3
bcsG	Z4954	Predicted inner membrane protein	2.09E-03	4.55E-03	-1.3
bioD	Z0997	Dethiobiotin synthetase	3.28E-04	1.84E-03	-1.3
blc	Z5756	Outer membrane lipoprotein (lipocalin)	3.45E-04	1.88E-03	-1.4
bssS	Z1697	Regulator of biofilm formation	8.85E-05	9.15E-04	-1.6
cbpA	Z1418	Curved DNA-binding protein, DnaJ homologue that functions as a co-chaperone of DnaK	1.18E-03	3.50E-03	-1.5
cfa	Z2686	Cyclopropane fatty acyl phospholipid synthase	1.88E-05	4.25E-04	-2.6
cheW	Z2941	Purine-binding chemotaxis protein	3.74E-04	1.96E-03	-1.3
coaE	Z0113	Putative DNA repair protein	5.84E-04	2.43E-03	-1.4
copA	Z0604	Copper transporter	9.69E-04	3.16E-03	-1.3
coxT	Z2970	Putative regulator for prophage CP-933T	6.06E-06	2.06E-04	-1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
creC	Z6002	Sensory histidine kinase in two-component regulatory system with CreB or PhoB, regulator of the CreBC regulon	5.60E-03	7.37E-03	-1.3
crr	Z3683	Glucose-specific enzyme IIA component of PTS	2.99E-06	1.52E-04	-2.0
csgA	Z1676	Cryptic curlin major subunit	2.87E-04	1.71E-03	-1.3
cydB	Z0901	Cytochrome d terminal oxidase, subunit II	1.23E-03	3.60E-03	-1.5
cysB	Z2535	DNA-binding transcriptional dual regulator, O-acetyl-L-serine-binding	1.76E-03	4.28E-03	-1.3
cysD	Z4060	Sulfate adenylyltransferase, subunit 2	1.55E-05	3.73E-04	-1.5
cysE	Z5034	Serine acetyltransferase	1.10E-03	3.39E-03	-1.3
cysJ	Z4074	Sulfite reductase, alpha subunit, flavoprotein	1.15E-04	1.04E-03	-1.6
cysK	Z3680	Cysteine synthase A, O-acetylserine sulfhydrolase A subunit	3.56E-06	1.63E-04	-3.7
cysN	Z4059	Sulfate adenylyltransferase, subunit 1	1.16E-03	3.49E-03	-1.5
ddpB	Z2224	Membrane component of an ABC superfamily D-ala-D-ala transporter	4.90E-03	6.94E-03	-1.3
ddpD	Z2226	D-ala-D-ala transporter subunit	5.19E-03	7.12E-03	-1.3
deoA	Z5984	Thymidine phosphorylase	8.37E-04	2.98E-03	-1.3
deoB	Z5985	Phosphopentomutase	3.96E-04	2.03E-03	-1.4
deoC	Z5983	2-deoxyribose-5-phosphate aldolase, NAD(P)-linked	2.38E-03	4.82E-03	-1.5
dkgA	Z4365	2,5-diketo-D-gluconate reductase A	2.41E-05	4.37E-04	-3.4
dnaC	Z5961	DNA biosynthesis protein	1.51E-02	1.30E-02	-1.3
dnaJ	Z0015	Chaperone Hsp40, co-chaperone with DnaK	5.38E-05	6.75E-04	-1.4
dnaK	Z0014	Chaperone Hsp70, co-chaperone with DnaJ	1.13E-08	2.21E-06	-7.1
*dos*_2	Z2221	C-terminal fragment of cAMP phosphodiesterase (pseudogene)	3.66E-03	5.94E-03	-1.3
dps	Z1034	Fe-binding and storage protein	5.42E-08	8.26E-06	-4.6
elaB	Z3526	Conserved protein	3.87E-06	1.71E-04	-2.2
eno	Z4094	Enolase	1.16E-03	3.49E-03	-2.1
epaO	Z4190	Type III secretion apparatus protein	2.04E-03	4.51E-03	-1.3
eutD	Z3714	Predicted phosphotransacetylase subunit	8.92E-04	3.06E-03	-1.3
fbaA	Z4263	Fructose-bisphosphate aldolase, class II	1.14E-06	8.65E-05	-4.8

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
fcl	Z3197	Fucose synthetase	4.78E-03	6.82E-03	-1.3
fdhF	Z5678	Formate dehydrogenase-H, selenopolypeptide subunit	2.90E-04	1.72E-03	-1.3
fhlA	Z4040	DNA-binding transcriptional activator	3.72E-03	5.99E-03	-1.3
fixC	Z0049	Predicted oxidoreductase with FAD/NAD(P)-binding domain	6.51E-03	7.99E-03	-1.3
flgB	Z1711	Flagellar component of cell-proximal portion of basal-body rod	1.58E-03	4.12E-03	-1.3
flgF	Z1715	Flagellar component of cell-proximal portion of basal-body rod	1.43E-04	1.18E-03	-1.3
flhD	Z2946	DNA-binding transcriptional dual regulator with FlhC	1.66E-03	4.22E-03	-1.3
flk	Z3583	Predicted flagella assembly protein	1.66E-03	4.22E-03	-1.3
frdA	Z5762	Fumarate reductase (anaerobic) catalytic and NAD/flavoprotein subunit	2.97E-05	4.67E-04	-1.6
frdB	Z5760	Fumarate reductase (anaerobic), Fe-S subunit	6.61E-03	8.06E-03	-1.3
frdC	Z5759	Fumarate reductase (anaerobic), membrane anchor subunit	3.38E-04	1.88E-03	-1.4
frdD	Z5758	Fumarate reductase (anaerobic), membrane anchor subunit	4.13E-04	2.08E-03	-1.4
frlC	Z4733	Predicted isomerase	3.44E-02	2.26E-02	-1.3
ftn	Z2960	Ferritin iron storage protein (cytoplasmic)	3.69E-04	1.95E-03	-1.4
gabD	Z3959	Succinate-semialdehyde dehydrogenase I, NADP-dependent	3.19E-04	1.81E-03	-1.3
gabT	Z3960	4-aminobutyrate aminotransferase, PLP-dependent	1.73E-03	4.27E-03	-1.3
gadA	Z4930	Glutamate decarboxylase A, PLP-dependent	2.48E-10	1.34E-07	-41.5
gadE	Z4925	Acid-induced positive regulator of glutamate-dependent acid resistance	1.03E-02	1.02E-02	-1.5
gadW	Z4928	DNA-binding transcriptional activator	1.50E-04	1.21E-03	-1.8
gadX	Z4929	DNA-binding transcriptional dual regulator	3.77E-04	1.97E-03	-2.1
galU	Z2012	Glucose-1-phosphate uridylyltransferase	9.11E-04	3.07E-03	-1.3
gapA	Z2818	Glyceraldehyde-3-phosphate dehydrogenase A	1.79E-03	4.28E-03	-1.5
gapC	Z2304	Glyceraldehyde-3-phosphate dehydrogenase	4.58E-06	1.90E-04	-3.2
gcvH	Z4241	Glycine cleavage complex lipoylprotein	1.41E-04	1.17E-03	-1.5
gcvP	Z4240	Glycine decarboxylase	2.76E-05	4.51E-04	-1.6
glgA	Z4791	Glycogen synthase	6.01E-04	2.44E-03	-2.0
glgP	Z4790	Glycogen phosphorylase	8.87E-04	3.06E-03	-2.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
gmd	Z3198	GDP-mannose dehydratase	1.32E-03	3.76E-03	-1.3
gntR	Z4806	DNA-binding transcriptional repressor	1.71E-03	4.27E-03	-1.3
gor	Z4900	Glutathione oxidoreductase	2.43E-03	4.89E-03	-1.3
gpmA	Z0925	Phosphoglyceromutase 1	3.43E-03	5.71E-03	-1.3
gpsA	Z5035	Glycerol-3-phosphate dehydrogenase (NAD+)	1.17E-05	3.22E-04	-1.7
groL	Z5748	Cpn60 chaperonin GroEL, large subunit of GroESL	7.78E-04	2.83E-03	-1.6
groS	Z5747	Cpn10 chaperonin GroES, small subunit of GroESL	3.10E-04	1.78E-03	-1.6
grpE	Z3907	Heat shock protein	8.49E-03	9.21E-03	-1.3
grxB	Z1701	Glutaredoxin 2 (Grx2)	5.45E-06	2.02E-04	-3.1
gshB	Z4292	Glutathione synthetase	4.18E-04	2.10E-03	-1.3
hdeA	Z4922	Stress response protein acid-resistance protein	4.80E-06	1.94E-04	-13.3
hflK	Z5781	Modulator for HflB protease specific for phage lambda cII repressor	1.41E-02	1.24E-02	-1.3
hfq	Z5779	Host factor I for bacteriophage Q beta replication, a growth-related protein	7.67E-05	8.39E-04	-1.7
holD	Z5973	DNA polymerase III, psi subunit	5.52E-03	7.30E-03	-1.3
hslU	Z5478	Molecular chaperone and ATPase component of HslUV protease	5.20E-05	6.60E-04	-1.6
hslV	Z5479	Peptidase component of the HslUV protease	6.08E-05	7.25E-04	-2.1
htpG	Z0590	Molecular chaperone HSP90 family	2.55E-05	4.37E-04	-1.3
hyaA	Z1389	Hydrogenase 1, small subunit	9.01E-03	9.46E-03	-1.4
hyaB	Z1390	Hydrogenase 1, large subunit	6.24E-06	2.06E-04	-2.2
hyaC	Z1391	Hydrogenase 1, b-type cytochrome subunit	4.21E-03	6.44E-03	-1.3
hyaF	Z1394	Protein involved in nickel incorporation into hydrogenase-1 proteins	9.87E-05	9.60E-04	-1.5
hycF	Z4028	Formate hydrogenlyase complex iron-sulfur protein	6.79E-04	2.62E-03	-1.4
hycG	Z4027	Hydrogenase 3 and formate hydrogenase complex, HycG subunit	4.75E-02	2.83E-02	-1.3
ivy	Z0277	Inhibitor of vertebrate C-lysozyme	3.62E-04	1.94E-03	-1.3
L7019	L7019	Transposase	3.61E-02	2.33E-02	-1.3
L7071	L7071	Unknown	1.11E-02	1.07E-02	-1.4
L7085	L7085	Hypothetical protein	8.52E-04	3.01E-03	-1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ligT	Z0158	2'-5' RNA ligase	9.34E-04	3.11E-03	-1.3
lysU	Z5732	Lysine tRNA synthetase, inducible	9.22E-05	9.37E-04	-2.3
map	Z0178	Methionine aminopeptidase	8.31E-03	9.09E-03	-1.3
mdoG	Z1683	Glucan biosynthesis protein, periplasmic	1.00E-04	9.62E-04	-1.5
mdoH	Z1684	Glucan biosynthesis: glycosyl transferase	2.35E-03	4.82E-03	-1.3
mdtD	Z3246	Multidrug efflux system protein	2.37E-04	1.55E-03	-1.3
metE	Z5351	5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	9.93E-04	3.19E-03	-1.4
metI	Z0210	D- and L-methionine transport protein (ABC superfamily, membrane)	2.33E-05	4.37E-04	-2.1
metK	Z4287	Methionine adenosyltransferase 1	9.77E-06	2.79E-04	-2.3
metN	Z0211	D- and L-methionine transport protein	3.53E-04	1.91E-03	-1.7
metQ	Z0209	D-methionine transport protein	2.36E-05	4.37E-04	-1.7
miaA	Z5778	Delta(2)-isopentenylpyrophosphate tRNA-adenosine transferase	1.61E-04	1.26E-03	-1.8
moaD	Z1003	Molybdopterin synthase, small subunit	9.20E-03	9.61E-03	-1.3
msyB	Z1686	Predicted protein	3.99E-05	5.70E-04	-1.9
mtfA	Z3132	Conserved protein	3.86E-05	5.69E-04	-1.4
mtlD	Z5024	Mannitol-1-phosphate dehydrogenase, NAD(P)-binding	4.66E-05	6.14E-04	-2.3
mviM	Z1705	Predicted oxidoreductase with NAD(P)-binding Rossmann-fold domain	3.90E-05	5.69E-04	-1.3
nadC	Z0119	Quinolinate phosphoribosyltransferase	2.43E-05	4.37E-04	-1.3
narG	Z2001	Nitrate reductase 1, alpha subunit	1.25E-04	1.10E-03	-1.5
narJ	Z2003	Molybdenum-cofactor-assembly chaperone subunit	7.50E-04	2.79E-03	-1.4
narW	Z2246	Nitrate reductase 2 (NRZ), delta subunit (assembly subunit)	1.77E-03	4.28E-03	-1.3
*nhoA*_1	Z2250	N-terminal fragment of N-hydroxyarylamine O-acetyltransferase	2.43E-05	4.37E-04	-1.5
*nhoA*_2	Z2249	C-terminal fragment of N-hydroxyarylamine O-acetyltransferase	2.21E-04	1.53E-03	-1.6
*nleG*2-2	Z2339	Non-LEE-encoded type III secreted effector	9.43E-04	3.13E-03	-1.4
*nleG*5-1	Z2337	Non-LEE-encoded type III secreted effector	5.32E-04	2.35E-03	-1.3
*nleH*1-1	Z0989	Non-LEE-encoded type III secreted effector	3.18E-03	5.49E-03	-1.3
nlpD	Z4050	Predicted outer membrane lipoprotein	5.30E-06	2.02E-04	-1.9

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
nlpE	Z0204	Lipoprotein involved with copper homeostasis and adhesion	2.05E-03	4.53E-03	-1.3
nrdH	Z3975	Glutaredoxin-like protein	1.56E-03	4.11E-03	-1.3
nuoB	Z3546	NADH:ubiquinone oxidoreductase, chain B	6.41E-03	7.92E-03	-1.3
nuoC	Z3545	NADH:ubiquinone oxidoreductase, chain C,D	1.02E-06	8.65E-05	-1.5
nuoE	Z3544	NADH:ubiquinone oxidoreductase, chain E	2.05E-05	4.25E-04	-1.8
nuoH	Z3541	NADH:ubiquinone oxidoreductase, membrane subunit H	6.81E-04	2.62E-03	-1.4
nuoJ	Z3539	NADH:ubiquinone oxidoreductase, membrane subunit J	2.88E-03	5.24E-03	-1.3
nuoL	Z3537	NADH:ubiquinone oxidoreductase, membrane subunit L	5.31E-03	7.18E-03	-1.3
nuoN	Z3534	NADH:ubiquinone oxidoreductase, membrane subunit N	1.93E-05	4.25E-04	-1.3
ompA	Z1307	Outer membrane protein A	4.36E-08	7.48E-06	-2.5
ompX	Z1036	Outer membrane protein	4.66E-07	5.41E-05	-1.3
oppA	Z2019	Oligopeptidase transporter subunit superfamily	2.45E-04	1.57E-03	-1.6
otsA	Z2949	Trehalose-6-phosphate synthase	2.16E-05	4.36E-04	-1.8
otsB	Z2950	Trehalose-6-phosphate phosphatase, biosynthetic	2.20E-06	1.20E-04	-1.4
pcm	Z4051	L-isoaspartate protein carboxylmethyltransferase type II	4.53E-04	2.18E-03	-1.3
pcnB	Z0154	Poly(A) polymerase I	1.54E-05	3.73E-04	-1.5
pdxK	Z3684	Pyridoxal-pyridoxamine kinase/hydroxymethylpyrimidine kinase	6.31E-06	2.06E-04	-1.6
pepD	Z0298	Aminoacyl-histidine dipeptidase (peptidase D)	2.50E-05	4.37E-04	-1.7
pfkB	Z2752	6-phosphofructokinase II	5.24E-04	2.34E-03	-1.4
pgk	Z4265	Phosphoglycerate kinase	6.92E-04	2.65E-03	-1.4
pgl	Z0938	6-phosphogluconolactonase	2.23E-04	1.53E-03	-1.6
pqqL	Z2214	Predicted peptidase	1.92E-05	4.25E-04	-1.5
prfC	Z5976	Peptide chain release factor RF-3	2.19E-03	4.68E-03	-1.3
proS	Z0206	Prolyl-tRNA synthetase	3.37E-03	5.68E-03	-1.3
psiA	L7088	Hypothetical protein	1.87E-03	4.34E-03	-1.5
pta	Z3559	Phosphate acetyltransferase	6.21E-06	2.06E-04	-1.6
ptsI	Z3682	PEP-protein phosphotransferase of PTS system (enzyme I)	1.09E-04	1.03E-03	-1.7

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ravA	Z5247	AAA+ family ATPase	5.19E-03	7.12E-03	-1.3
rcsF	Z0208	Predicted outer membrane protein, signal	5.40E-04	2.35E-03	-1.4
rfaD	Z5046	ADP-L-glycero-D-mannoheptose-6-epimerase, NAD(P)-binding	1.36E-05	3.52E-04	-1.7
rfaF	Z5047	ADP-heptose:LPS heptosyltransferase II	7.16E-03	8.34E-03	-1.3
rffH	Z5300	Glucose-1-phosphate thymidylyltransferase	8.68E-05	9.09E-04	-1.3
rimI	Z5974	Acetylase for 30S ribosomal subunit protein S18	5.24E-03	7.16E-03	-1.3
rnr	Z5786	Exoribonuclease R, RNase R	5.14E-04	2.33E-03	-1.3
rorf1	Z5143	LEE-encoded type III secretion system factor	2.82E-03	5.21E-03	-1.3
rpoS	Z4049	RNA polymerase, sigma S (sigma 38) factor	3.12E-06	1.53E-04	-2.1
rsmC	Z5972	16S rRNA m2G1207 methylase	1.36E-03	3.78E-03	-1.3
sapA	Z2494	Predicted antimicrobial peptide transporter subunit	2.59E-03	5.02E-03	-1.3
sapF	Z2500	Peptide transport system ATP-binding protein	1.94E-02	1.53E-02	-1.3
sdhB	Z0878	Succinate dehydrogenase, FeS subunit	4.47E-03	6.56E-03	-1.3
secB	Z5036	Protein export chaperone	3.84E-03	6.07E-03	-1.3
serA	Z4251	D-3-phosphoglycerate dehydrogenase	5.43E-04	2.35E-03	-1.3
slp	Z4908	Outer membrane lipoprotein	4.26E-03	6.47E-03	-1.6
speD	Z0130	S-adenosylmethionine decarboxylase	1.29E-05	3.46E-04	-1.7
speE	Z0131	Spermidine synthase (putrescine aminopropyltransferase)	3.07E-04	1.78E-03	-1.4
sucA	Z0880	2-oxoglutarate decarboxylase, thiamin-requiring	1.02E-04	9.62E-04	-1.5
sucD	Z0883	Succinyl-CoA synthetase, NAD(P)-binding, alpha subunit	2.55E-05	4.37E-04	-1.3
sufB	Z2711	Component of SufBCD complex	1.95E-05	4.25E-04	-2.4
sufC	Z2710	Transport protein associated with Fe-S cluster assembly	4.14E-05	5.79E-04	-1.6
sufE	Z2707	Sulfur acceptor protein	2.64E-05	4.37E-04	-1.4
sxy	Z1309	CRP-S promoter co-activator	1.12E-03	3.43E-03	-1.3
talA	Z3720	Transaldolase A	1.93E-06	1.20E-04	-4.4
talB	Z0008	Transaldolase B	1.80E-03	4.28E-03	-1.3
tdcE	Z4466	Pyruvate formate-lyase 4/2-ketobutyrate formate-lyase	7.35E-05	8.20E-04	-1.7

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
tdh	Z5043	Threonine 3-dehydrogenase, NAD(P)-binding	1.54E-02	1.31E-02	-1.3
tktA	Z4279	Transketolase 1, thiamin-binding	5.65E-04	2.39E-03	-1.5
tktB	Z3721	Transketolase 2, thiamin-binding	1.12E-03	3.43E-03	-1.4
tolC	Z4392	Transport channel	4.54E-06	1.90E-04	-2.0
toxB	L7095	Putative cytotoxin	1.94E-02	1.53E-02	-1.3
treF	Z4932	Cytoplasmic trehalase	4.20E-02	2.61E-02	-1.3
trpR	Z5995	DNA-binding transcriptional repressor, tryptophan-binding	1.48E-03	4.00E-03	-1.3
truA	Z3580	tRNA pseudouridine synthase A	2.75E-04	1.67E-03	-1.3
ushA	Z0599	Bifunctional UDP-sugar hydrolase and 5'-nucleotidase	1.16E-04	1.04E-03	-1.6
waaD	Z5050	UDP-glucose:(glucosyl) LPS alpha-1,2-glucosyltransferase	1.21E-03	3.57E-03	-1.4
waaJ	Z5051	UDP-glucose:(galactosyl) LPS alpha-1,2-glucosyltransferase	8.08E-03	8.97E-03	-1.3
waaL	Z5049	Putative LPS biosynthesis enzyme	6.67E-05	7.78E-04	-1.4
wbdP	Z3199	Glycosyl transferase	4.96E-04	2.27E-03	-1.4
wbdQ	Z3196	GDP-mannose mannosylhydrolase	2.88E-03	5.24E-03	-1.3
yacC	Z0132	Predicted protein	5.29E-04	2.35E-03	-1.4
yacG	Z0111	Zinc-binding protein	5.71E-03	7.45E-03	-1.3
yaeB	Z0207	Conserved protein	6.05E-03	7.66E-03	-1.4
yaeF	Z0205	Predicted lipoprotein	2.46E-03	4.89E-03	-1.3
yaeQ	Z0202	Conserved protein (pseudogene)	4.58E-04	2.18E-03	-1.6
yafD	Z0232	Endonuclease-exonuclease-phosphatase family member	3.92E-04	2.02E-03	-1.5
yafO	Z0294	Predicted toxin of the YafO-YafN toxin-antitoxin system	1.05E-02	1.03E-02	-1.3
yagU	Z0353	Conserved inner membrane protein	4.83E-05	6.25E-04	-1.3
ybaS	Z0606	Predicted glutaminase	1.99E-04	1.45E-03	-1.8
ybaT	Z0607	Predicted transporter	1.40E-06	9.63E-05	-1.7
ybgA	Z0858	Conserved protein	1.77E-05	4.16E-04	-1.8
ybgH	Z0860	Predicted transporter	3.16E-03	5.48E-03	-1.3
ybgL	Z0864	Predicted lactam utilization protein	4.97E-04	2.27E-03	-1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ybhF	Z1014	Fused predicted transporter subunits of ABC superfamily	2.71E-03	5.12E-03	-1.6
ybhG	Z1015	Predicted membrane fusion protein (MFP) component of efflux pump, membrane anchor	1.28E-03	3.70E-03	-1.5
ybhL	Z1005	Predicted inner membrane protein	8.73E-03	9.29E-03	-1.3
ybhO	Z1008	Cardiolipin synthase 2	1.61E-02	1.35E-02	-1.3
ybiB	Z1021	Predicted transferase/phosphorylase	1.58E-03	4.12E-03	-1.4
ybiC	Z1022	Predicted dehydrogenase	2.73E-04	1.66E-03	-1.3
ybiJ	Z1023	Predicted protein	1.32E-04	1.15E-03	-1.3
yccA	Z1322	Inner membrane protein	8.61E-04	3.02E-03	-1.3
yccS	Z1311	Predicted inner membrane protein	5.53E-03	7.31E-03	-1.3
yceK	Z1685	Predicted lipoprotein	9.17E-04	3.08E-03	-1.3
ychP	Z1995	Predicted invasin	1.57E-03	4.12E-03	-1.3
yddE	Z2248	Conserved protein	1.49E-02	1.29E-02	-1.3
ydiA	Z2732	Conserved protein	2.79E-04	1.69E-03	-1.4
ydiT	Z2729	Predicted 4Fe-4S ferredoxin-type protein	3.38E-03	5.68E-03	-1.3
ydiZ	Z2753	Predicted protein	1.79E-05	4.16E-04	-1.6
yeaG	Z2823	Conserved protein with nucleoside triphosphate hydrolase domain	5.13E-07	5.41E-05	-3.4
yeaH	Z2824	Conserved protein	1.12E-06	8.65E-05	-3.4
yebW	Z2884	Predicted protein	6.67E-04	2.58E-03	-1.3
yecA	Z2964	Conserved metal-binding protein	1.54E-04	1.23E-03	-1.3
yehB	Z3277	Predicted outer membrane protein	4.27E-02	2.64E-02	-1.3
yfbB	Z3523	Predicted peptidase	2.22E-03	4.72E-03	-1.3
yfbQ	Z3551	Predicted aminotransferase	1.72E-02	1.41E-02	-1.3
yfbT	Z3554	Sugar phosphatase	3.26E-03	5.56E-03	-1.3
yfeT	Z3692	Predicted DNA-binding transcriptional regulator	4.52E-04	2.18E-03	-1.3
yfeZ	Z3698	Predicted inner membrane protein	6.58E-04	2.56E-03	-1.3
ygaM	Z3972	Predicted protein	4.77E-05	6.23E-04	-1.9

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ygbE	Z4057	Conserved inner membrane protein	5.05E-03	7.03E-03	-1.3
ygdH	Z4112	Conserved protein	3.15E-03	5.48E-03	-1.6
yggE	Z4259	Conserved protein	2.62E-05	4.37E-04	-1.7
yggJ	Z4291	Predicted protein	1.05E-02	1.03E-02	-1.3
yggR	Z4295	Predicted transporter	2.58E-04	1.60E-03	-1.4
ygiC	Z4395	Predicted enzyme	6.98E-03	8.25E-03	-1.3
ygiW	Z4376	Conserved protein	1.37E-06	9.63E-05	-1.9
ygjG	Z4426	Putrescine:2-oxoglutaric acid aminotransferase	5.41E-03	7.24E-03	-1.3
yhaH	Z4457	Predicted inner membrane protein	1.13E-04	1.04E-03	-1.9
yhbG	Z4564	Predicted transporter subunit: ATP-binding component of ABC superfamily	5.85E-04	2.43E-03	-1.3
yhdA	Z4611	Hypothetical protein	1.24E-02	1.15E-02	-1.3
yheL	Z4701	Subunit of the heterohexameric YheLMN (TusBCD) mediator of sulfur relay	6.74E-03	8.11E-03	-1.3
yheO	Z4704	Predicted DNA-binding transcriptional regulator	5.22E-04	2.34E-03	-1.5
yhiM	Z4890	Conserved inner membrane protein	1.40E-05	3.54E-04	-2.4
yicL	Z5146	Predicted inner membrane protein	1.89E-04	1.40E-03	-1.3
yicS	Z5148	Predicted protein	8.52E-03	9.21E-03	-1.3
yieP	Z5258	Predicted transcriptional regulator	2.41E-04	1.56E-03	-1.3
yiiT	Z5468	Stress-induced protein	1.24E-02	1.15E-02	-1.6
yjbQ	Z5655	Conserved protein	5.41E-05	6.75E-04	-1.8
yjdI	Z5728	Conserved protein	2.27E-05	4.37E-04	-2.0
yjdJ	Z5729	Predicted acyltransferase with acyl-CoA N-acyltransferase domain	3.00E-05	4.67E-04	-1.7
yjeT	Z5783	Conserved inner membrane protein	5.39E-04	2.35E-03	-1.6
ykgC	Z0381	Predicted oxidoreductase with FAD/NAD(P)-binding domain and dimerization domain	2.57E-04	1.60E-03	-1.4
ykgD	Z0382	Predicted DNA-binding transcriptional regulator	1.32E-03	3.76E-03	-1.3
ymdB	Z1679	Conserved protein	5.09E-06	1.99E-04	-1.4
ymdC	Z1680	Predicted hydrolase	1.91E-03	4.38E-03	-1.4

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ynbA	Z2319	Predicted inner membrane protein	1.48E-05	3.68E-04	-1.7
ypfJ	Z3734	Conserved protein	8.58E-03	9.24E-03	-1.3
yqcE	Z4086	Predicted transporter	5.65E-04	2.39E-03	-1.3
yqhD	Z4364	NADP-dependent alcohol dehydrogenase	1.31E-03	3.75E-03	-1.5
yqjC	Z4451	Conserved protein	5.07E-07	5.41E-05	-2.5
yqjD	Z4452	Conserved protein	7.55E-03	8.58E-03	-1.7
yqjE	Z4453	Conserved inner membrane protein	3.07E-04	1.78E-03	-2.4
yqjI	Z4424	Predicted transcriptional regulator	4.51E-03	6.58E-03	-1.3
Z0260	Z0260	Predicted protein	1.14E-04	1.04E-03	-1.3
Z0262	Z0262	Predicted protein	4.78E-04	2.23E-03	-1.3
Z0264	Z0264	Predicted protein	3.76E-03	6.02E-03	-1.3
Z0609	Z0609	Predicted protein	3.87E-04	2.02E-03	-1.4
Z0615	Z0615	Putative RTX family exoprotein	6.49E-05	7.67E-04	-1.5
Z0634	Z0634	Putative cytoplasmic membrane export protein	7.71E-05	8.39E-04	-1.6
Z0705	Z0705	Predicted protein	2.23E-03	4.72E-03	-1.3
Z0706	Z0706	Predicted protein	2.83E-03	5.21E-03	-1.3
Z0885	Z0885	Putative LysR-like transcriptional regulator	1.61E-03	4.15E-03	-1.3
Z0892	Z0892	Putative methylaspartate ammonia-lyase	1.04E-03	3.29E-03	-1.3
Z0894	Z0894	Predicted protein	3.41E-04	1.88E-03	-1.3
Z0895	Z0895	Putative glutamate mutase subunit S	2.98E-04	1.75E-03	-1.3
Z0896	Z0896	Predicted protein	3.64E-02	2.34E-02	-1.3
Z0980	Z0980	Putative tail component of prophage CP-933K	1.30E-02	1.19E-02	-1.3
Z1335	Z1335	Unknown protein encoded by cryptic prophage CP-933M	1.46E-02	1.27E-02	-1.3
Z1340	Z1340	Unknown protein encoded by cryptic prophage CP-933M	7.34E-03	8.47E-03	-1.3
Z1444	Z1444	Putative serine/threonine kinase encoded by bacteriophage BP-933W	6.80E-05	7.81E-04	-1.5
Z1456	Z1456	Unknown protein encoded by bacteriophage BP-933W	2.20E-03	4.69E-03	-1.3
Z1498	Z1498	Unknown protein encoded by bacteriophage BP-933W	2.13E-03	4.59E-03	-1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
Z1844	Z1844	Unknown protein encoded by prophage CP-933C	9.02E-04	3.07E-03	-1.3
Z2323	Z2323	Internal fragment of a predicted autotransporter (pseudogene)	9.24E-09	2.11E-06	-3.0
Z2393	Z2393	Unknown protein encoded within prophage CP-933R	5.98E-04	2.44E-03	-1.3
Z2400	Z2400	Partial putative repressor protein encoded within prophage CP-933R	1.03E-03	3.28E-03	-1.6
Z2403	Z2403		2.16E-03	4.65E-03	-1.3
Z2435	Z2435	Stress-induced protein	2.17E-06	1.20E-04	-3.0
Z2506	Z2506	N-terminal fragment of a putative outer membrane channel protein	4.37E-03	6.50E-03	-1.3
Z2967	Z2967	Unknown protein encoded by prophage CP-933T	1.41E-02	1.24E-02	-1.3
Z2971	Z2971	Unknown protein encoded by prophage CP-933T	3.57E-04	1.92E-03	-1.3
Z2973	Z2973	Unknown protein encoded by prophage CP-933T	9.06E-04	3.07E-03	-1.3
Z2974	Z2974	Unknown protein encoded by prophage CP-933T	7.30E-03	8.45E-03	-1.3
Z2975	Z2975	Unknown protein encoded by prophage CP-933T	4.43E-03	6.55E-03	-1.3
Z2976	Z2976	Unknown protein encoded by prophage CP-933T	2.24E-04	1.53E-03	-1.3
Z2978	Z2978	Putative replication protein for prophage CP-933T	1.73E-03	4.27E-03	-1.3
Z3082	Z3082	Putative tail fiber component L of prophage CP-933U	1.18E-04	1.05E-03	-1.3
Z3122	Z3122	Unknown protein encoded within prophage CP-933U	1.60E-04	1.25E-03	-1.3
Z3126	Z3126	Putative repressor protein of prophage CP-933U	1.05E-03	3.31E-03	-1.3
Z3332	Z3332	Unknown protein encoded within prophage CP-933V	4.32E-04	2.12E-03	-1.4
Z3939	Z3939	Predicted protein	5.68E-04	2.39E-03	-1.5
Z4192	Z4192	Hypothetical protein	1.62E-02	1.35E-02	-1.3
Z4267	Z4267	Hypothetical protein	4.95E-04	2.27E-03	-1.3
Z4852	Z4852	Putative phospholipid biosynthesis acyltransferase	2.48E-03	4.91E-03	-1.3
Z5898	Z5898	Predicted protein	9.57E-03	9.82E-03	-1.4
Up-regulated					
aat	Z1229	Leucyl/phenylalanyl-tRNA-protein transferase	9.40E-05	0.0009	2.0
ada	Z3471	Fused DNA-binding transcriptional dual regulator	8.60E-03	0.0092	1.4
agaV	Z4485	N-acetylgalactosamine-specific enzyme IIB component of PTS	1.36E-02	0.0122	1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ahpC	Z0749	Alkyl hydroperoxide reductase, C22 subunit	2.60E-03	0.0050	1.9
ahpF	Z0750	Alkyl hydroperoxide reductase, F52a subunit, FAD/NAD(P)-binding	1.49E-03	0.0040	2.0
ais	Z3510	Conserved protein	2.65E-02	0.0188	1.3
alkA	Z3237	3-methyl-adenine DNA glycosylase II	8.49E-04	0.0030	1.5
alkB	Z3470	Oxidative demethylase of N1-methyladenine or N3-methylcytosine DNA lesions	4.46E-02	0.0271	1.5
anmK	Z2654	Anhydro-N-acetylmuramic acid kinase	2.75E-02	0.0193	1.3
apaG	Z0059	Protein associated with Co2+ and Mg2+ efflux	6.30E-03	0.0078	1.6
apaH	Z0058	Diadenosine tetraphosphatase	2.75E-02	0.0193	1.5
aqpZ	Z1109	Aquaporin	8.35E-03	0.0091	1.8
araA	Z0070	L-arabinose isomerase	2.01E-03	0.0045	1.7
araB	Z0072	L-ribulokinase	1.18E-03	0.0035	1.6
araC	Z0073	DNA-binding transcriptional dual regulator	8.08E-03	0.0090	1.5
araD	Z0069	L-ribulose-5-phosphate 4-epimerase	2.68E-03	0.0051	1.5
argB	Z5517	Acetylglutamate kinase	2.08E-03	0.0045	1.5
argE	Z5515	Acetylornithine deacetylase	4.36E-04	0.0021	1.3
argH	Z5518	Argininosuccinate lyase	1.01E-02	0.0101	1.3
arnC	Z3512	Undecaprenyl-phosphate 4-amino-4-deoxy-L-arabinose transferase	9.99E-03	0.0101	1.3
aroA	Z1254	3-phosphoshikimate 1-carboxyvinyltransferase	1.78E-03	0.0043	1.5
aroB	Z4742	3-dehydroquinate synthase	1.26E-02	0.0116	1.3
aroK	Z4743	Shikimate kinase I	1.85E-02	0.0148	1.2
artI	Z1093	Arginine transporter subunit; periplasmic-binding component of ABC superfamily	8.68E-03	0.0093	1.8
artJ	Z1090	Arginine 3rd transport system periplasmic binding protein	2.37E-04	0.0015	1.7
artM	Z1091	Arginine transporter subunit; membrane component of ABC superfamily	3.07E-03	0.0054	1.7
artP	Z1094	Arginine transporter subunit; ATP-binding component of ABC superfamily	1.70E-03	0.0043	1.8
artQ	Z1092	Arginine transporter subunit; membrane component of ABC superfamily	5.98E-04	0.0024	2.1
asmA	Z3232	Predicted assembly protein	1.96E-02	0.0153	1.4

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
asnS	Z1278	Asparaginyl tRNA synthetase	4.49E-04	0.0022	1.6
aspC	Z1275	Aspartate transaminase	2.71E-03	0.0051	1.6
astB	Z2777	Succinylarginine dihydrolase	1.38E-04	0.0012	1.3
bipA	Z5407	GTP-binding protein	5.09E-03	0.0071	1.4
borW	Z1474	Putative Bor protein precursor of bacteriophage BP-933W	6.16E-11	0.0000	2.9
ccmB	Z3457	Heme exporter subunit	5.93E-03	0.0076	1.6
ccmC	Z3456	Heme exporter subunit	7.78E-03	0.0088	1.6
ccmD	Z3455	Cytochrome c biogenesis protein	2.38E-02	0.0175	1.3
cho	Z2771	Endonuclease of nucleotide excision repair	2.37E-02	0.0174	1.3
citA	Z0764	Sensory histidine kinase in two-component regulatory system with *citB*	3.40E-03	0.0057	1.6
citB	Z0765	DNA-binding response regulator in two-component regulatory system with *citA*	2.84E-03	0.0052	1.7
citC	Z0762	Citrate lyase synthetase	1.30E-03	0.0037	1.7
citD	Z0761	Citrate lyase, acyl carrier (gamma) subunit	1.13E-02	0.0109	1.4
citE	Z0760	Citrate lyase, citryl-ACP lyase (beta) subunit	3.69E-04	0.0020	2.2
citF	Z0759	Citrate lyase, citrate-ACP transferase (alpha) subunit	2.98E-03	0.0054	1.8
citG	Z0757	Triphosphoribosyl-dephospho-CoA transferase	7.12E-03	0.0083	1.6
citT	Z0756	Citrate:succinate antiporter	2.61E-05	0.0004	1.9
citX	Z0758	Apo-citrate lyase phosphoribosyl-dephospho-CoA transferase	3.80E-03	0.0060	1.9
clpA	Z1119	ATPase and specificity subunit of ClpA-ClpP ATP-dependent serine protease	5.20E-04	0.0023	1.7
clpB	Z3886	Protein disaggregation chaperone	3.02E-02	0.0206	1.3
clpS	Z1118	Regulatory protein for ClpA substrate specificity	3.74E-03	0.0060	1.6
cmk	Z1256	Cytidylate kinase	4.07E-03	0.0063	1.8
crcA	Z0767	Palmitoyl transferase for Lipid A	6.14E-04	0.0025	2.1
crcB	Z0770	Predicted inner membrane protein associated with chromosome condensation	6.45E-03	0.0079	1.6
cscB	Z3623	Sucrose permease	3.65E-02	0.0235	1.3
csdA	Z4127	Cysteine sulfinate desulfinase	1.80E-02	0.0145	1.2

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
cspD	Z1117	Cold shock protein homolog	1.91E-03	0.0044	1.6
cstA	Z0740	Carbon starvation protein	4.23E-05	0.0006	2.2
cvrA	Z1954	Predicted cation/proton antiporter	9.70E-04	0.0032	1.7
cycA	Z5819	D-alanine/D-serine/glycine transporter	7.26E-03	0.0084	1.4
cydC	Z1230	ATP-binding component of cytochrome-related transport	9.04E-04	0.0031	1.8
cydD	Z1231	Cysteine transporter subunit of ABC superfamily	5.56E-04	0.0024	1.8
dacB	Z4544	D-alanyl-D-alanine carboxypeptidase	1.03E-02	0.0102	1.4
dadA	Z1952	D-amino acid dehydrogenase	7.73E-03	0.0087	1.9
dadX	Z1953	Alanine racemase 2, PLP-binding	3.53E-05	0.0005	1.7
dcp	Z2160	Dipeptidyl carboxypeptidase II	1.97E-02	0.0154	1.3
dcuC	Z0766	Anaerobic C4-dicarboxylate transport	1.36E-04	0.0012	2.1
ddlB	Z0102	D-alanine:D-alanine ligase	3.75E-03	0.0060	1.4
degQ	Z4593	Serine endoprotease, periplasmic	4.46E-02	0.0271	1.3
degS	Z4594	Serine endoprotease, periplasmic	1.41E-02	0.0124	1.4
dhaK	Z1971	Dihydroxyacetone kinase, N-terminal domain	1.78E-04	0.0014	2.1
dhaL	Z1970	Dihydroxyacetone kinase, C-terminal domain	5.97E-04	0.0024	2.2
dhaM	Z1969	Fused predicted dihydroxyacetone-specific PTS enzymes	6.08E-04	0.0025	1.9
dinD	Z5070	DNA-damage-inducible protein	1.69E-02	0.0139	1.3
dmsB	Z1241	Dimethyl sulfoxide reductase, anaerobic, subunit B	9.65E-03	0.0098	1.6
dmsC	Z1242	Dimethyl sulfoxide reductase, anaerobic, subunit C	7.86E-03	0.0088	1.5
dnaG	Z4419	DNA primase	6.54E-04	0.0026	1.4
dsbB	Z1948	Oxidoreductase that catalyzes reoxidation of DsbA protein disulfide isomerase I	5.00E-03	0.0070	1.7
dsbG	Z0748	Periplasmic disulfide isomerase/thiol-disulphide oxidase	4.00E-03	0.0062	2.0
dusB	Z4620	tRNA-dihydrouridine synthase B	1.76E-03	0.0043	1.3
eco	Z3467	Ecotin, a serine protease inhibitor	6.63E-03	0.0081	1.3
efeB	Z1521	Redox component of a tripartite ferrous iron transporter	0.0066005	0.008059	1.2
efeO	Z1520	Component of a tripartite ferrous iron transporter	4.48E-05	0.0006	1.7

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
efeU	Z1519	Ferrous iron permease	8.48E-04	0.0030	1.5
eivA	Z4195	Type III secretion apparatus protein	1.18E-02	0.0112	1.3
eivE	Z4196	Putative secreted protein	7.30E-03	0.0085	1.3
eivG	Z4197	Type III secretion apparatus protein	6.33E-04	0.0025	1.3
emtA	Z1956	Lytic murein endotransglycosylase E	3.00E-04	0.0018	1.8
entA	Z0738	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase	4.83E-04	0.0022	1.7
entB	Z0737	2,3-dihydro-2,3-dihydroxybenzoate synthetase, isochroismatase (Enterobactin synthetase component B)	1.18E-02	0.0112	1.6
entC	Z0735	Isochorismate synthase 1	2.17E-02	0.0165	1.5
entD	Z0723	Phosphopantetheinyltransferase component of enterobactin synthase multienzyme complex	5.71E-03	0.0075	1.7
entE	Z0736	Enterobactin synthetase component E	2.26E-04	0.0015	1.8
entF	Z0727	ATP-dependent serine activating enzyme (may be part of enterobactin synthase as component F)	1.28E-02	0.0118	1.7
era	Z3847	Membrane-associated, 16S rRNA-binding GTPase	1.98E-02	0.0154	1.3
escF	Z5103	LEE-encoded type III secretion system component	2.23E-04	0.0015	1.4
espA	Z5107	Secreted protein EspA	1.05E-02	0.0103	1.3
espD	Z5106	Secreted protein EspD	9.52E-03	0.0098	1.3
espF	Z5100	LEE-encoded type III secreted effector	3.78E-02	0.0241	1.3
espM1	Z2565	Non-LEE-encoded type III secreted effector	2.96E-02	0.0203	1.3
espR4	Z3026	Predicted non-LEE-encoded type III secreted effector	9.91E-04	0.0032	1.3
espY1	Z0065	Non-LEE-encoded type III secreted effector	3.09E-02	0.0209	1.4
espY2	Z0078	Predicted non-LEE-encoded type III secreted effector	4.28E-04	0.0021	1.7
exbD	Z4358	Membrane spanning protein in TonB-ExbB-ExbD complex	3.73E-02	0.0239	1.3
exoO	Z2037	Putative exonuclease VIII, ds DNA exonuclease, 5' --> 3' specific encoded by prophage CP-933O	4.14E-02	0.0258	1.3
fadR	Z1950	DNA-binding transcriptional dual regulator of fatty acid metabolism	3.30E-03	0.0056	1.9

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
fepA	Z0724	Iron-enterobactin outer membrane transporter	4.06E-04	0.0021	1.8
fepB	Z0734	Iron-enterobactin transporter subunit	1.32E-03	0.0038	2.0
fepC	Z0729	ATP-binding component of ferric enterobactin transport	2.62E-03	0.0050	1.8
fepD	Z0732	Ferric enterobactin (enterochelin) transport	8.50E-03	0.0092	1.8
fepE	Z0728	Regulator of length of O-antigen component of lipopolysaccharide chains	5.74E-03	0.0075	1.5
fepG	Z0731	Iron-enterobactin transporter subunit	1.03E-03	0.0033	1.8
fes	Z0725	Enterobactin/ferric enterobactin esterase	3.54E-02	0.0230	1.6
fis	Z4621	Global DNA-binding transcriptional dual regulator	2.09E-04	0.0015	1.6
fmlA	Z2200	Major subunit of F9 fimbriae	5.25E-03	0.0072	1.4
fmlB	Z2201	F9 fimbriae chaperone	2.36E-03	0.0048	1.3
focA	Z1250	Formate transporter	4.48E-03	0.0066	1.6
frlR	Z4736	Predicted DNA-binding transcriptional regulator	9.64E-03	0.0098	1.3
fruR	Z0090	DNA-binding transcriptional dual regulator	7.56E-03	0.0086	1.6
fsaB	Z5501	Fructose-6-phosphate aldolase 2	4.02E-03	0.0062	1.4
ftsA	Z0104	Cell division protein	4.35E-03	0.0065	1.4
ftsH	Z4540	Subunit of integral membrane ATP-dependent zinc metallopeptidase	2.31E-04	0.0015	1.5
ftsK	Z1235	DNA-binding membrane protein required for chromosome resolution and partitioning	4.70E-03	0.0068	1.7
ftsL	Z0093	Membrane bound cell division protein at septum containing leucine zipper motif	2.48E-03	0.0049	1.6
ftsQ	Z0103	Membrane anchored protein involved in growth of wall at septum	1.19E-02	0.0112	1.4
ftsW	Z0099	Cell division membrane protein	2.90E-03	0.0053	1.6
ftsZ	Z0105	GTP-binding tubulin-like cell division protein	2.41E-03	0.0049	1.5
fucP	Z4118	L-fucose transporter	1.27E-02	0.0117	1.2
fusA	Z4698	Protein chain elongation factor EF-G, GTP-binding	6.14E-03	0.0077	1.3
garP	Z4479	Predicted (D)-galactarate transporter	2.89E-02	0.0200	1.3
ghrB	Z4978	2-keto-D-gluconate reductase (glyoxalate reductase) (2-ketoaldonate reductase)	2.53E-02	0.0183	1.5

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
glpA	Z3499	sn-glycerol-3-phosphate dehydrogenase (anaerobic), large subunit, FAD/NAD(P)-binding	4.97E-03	0.0070	1.4
glpB	Z3500	sn-glycerol-3-phosphate dehydrogenase (anaerobic), membrane anchor subunit	3.47E-03	0.0058	1.5
glpC	Z3501	sn-glycerol-3-phosphate dehydrogenase (anaerobic), small subunit	4.87E-03	0.0069	1.5
glpD	Z4786	sn-glycerol-3-phosphate dehydrogenase, aerobic, FAD/NAD(P)-binding	6.84E-05	0.0008	2.1
glpQ	Z3497	Periplasmic glycerophosphodiester phosphodiesterase	8.68E-03	0.0093	1.6
glpT	Z3498	sn-glycerol-3-phosphate transporter	2.61E-03	0.0050	1.6
gpt	Z0299	Guanine-xanthine phosphoribosyltransferase	3.73E-04	0.0020	1.3
greA	Z4543	Transcription elongation factor	4.68E-02	0.0281	1.3
grxA	Z1076	Glutaredoxin 1, redox coenzyme for ribonucleotide reductase (RNR1a)	6.15E-03	0.0077	1.7
gyrA	Z3484	DNA gyrase (type II topoisomerase), subunit A	3.43E-02	0.0226	1.6
hcaF	Z3810	3-phenylpropionate dioxygenase, small (beta) subunit	2.69E-02	0.0190	1.3
hcp	Z1107	Hybrid cluster protein (HCP)	2.38E-03	0.0048	1.6
hcr	Z1106	HCP oxidoreductase, NADH-dependent	7.53E-04	0.0028	1.4
hemB	Z0468	Porphobilinogen synthase	3.34E-02	0.0221	1.3
hepA	Z0067	RNA polymerase-associated helicase protein	1.60E-03	0.0042	1.6
hipA	Z2197	Serine kinase (pseudogene)	3.03E-03	0.0054	1.3
hlyE	Z1944	Hemolysin E	7.32E-05	0.0008	1.7
hofQ	Z4744	Predicted fimbrial transporter	2.31E-02	0.0172	1.2
htrC	Z5563	Heat shock protein	2.93E-02	0.0202	1.3
hupA	Z5576	HU, DNA-binding transcriptional regulator, alpha subunit	1.52E-03	0.0041	1.3
ihfB	Z1258	Integration host factor (IHF), DNA-binding protein, beta subunit	2.01E-03	0.0045	2.0
ilvH	Z0088	Acetolactate synthase III, thiamin-dependent, small subunit	1.17E-03	0.0035	1.4
ilvL	Z5278	*ilvG* operon leader peptide	7.03E-04	0.0027	1.8
infA	Z1228	Translation initiation factor IF-1	9.18E-05	0.0009	2.1
infB	Z4529	Protein chain initiation factor 2, IF2	1.36E-05	0.0004	1.7
intO	Z2036	Putative integrase for prophage CP-933O	1.53E-02	0.0131	1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
iscR	Z3798	Fe-S cluster-containing transcription factor	4.11E-05	0.0006	1.6
kbaZ	Z4484	Tagatose-biphosphate aldolase 1, kbaZ subunit	9.55E-03	0.0098	1.4
kdsB	Z1264	3-deoxy-manno-octulosonate cytidylyltransferase	4.77E-03	0.0068	1.8
kefB	Z4710	Potassium:proton antiporter	1.15E-02	0.0110	1.3
kgtP	Z3872	Alpha-ketoglutarate transporter	3.93E-04	0.0020	1.5
ksgA	Z0060	S-adenosylmethionine-6-N',N'-adenosyl (rRNA) dimethyltransferase	8.68E-03	0.0093	1.5
L7072	L7072	Unknown	2.95E-02	0.0202	1.3
ldcA	Z1955	L,D-carboxypeptidase A	1.95E-03	0.0044	1.6
lepA	Z3851	GTP-binding membrane protein	1.08E-02	0.0105	1.3
leuA	Z0083	2-isopropylmalate synthase	2.61E-02	0.0187	1.5
leuC	Z0081	3-isopropylmalate isomerase subunit, dehydratase component	3.99E-02	0.0251	2.1
leuD	Z0080	3-isopropylmalate isomerase subunit	2.74E-02	0.0193	2.5
leuO	Z0086	DNA-binding transcriptional activator	1.28E-02	0.0118	1.5
lexA	Z5642	DNA-binding transcriptional repressor of SOS regulon	4.22E-04	0.0021	2.6
lhr	Z2673	Predicted ATP-dependent helicase	7.41E-03	0.0085	1.5
lipA	Z0773	Lipoate synthase	3.17E-03	0.0055	1.9
lipB	Z0775	Lipoyl-protein ligase	3.37E-04	0.0019	2.3
lolA	Z1237	Chaperone for lipoproteins	3.53E-02	0.0230	1.6
lomP	Z6028	Putative Lom-like outer membrane protein of cryptic prophage CP-933P	3.97E-03	0.0062	1.3
lpfD	Z4966	Putative fimbrial protein	1.71E-02	0.0141	1.3
lpfD2	Z5221	Putative fimbrial protein	2.58E-02	0.0185	1.3
lpfE	Z4965	Putative fimbrial subunit	1.31E-02	0.0119	1.3
lpp	Z2705	Murein lipoprotein	1.08E-03	0.0034	1.3
lpxC	Z0106	UDP-3-O-acyl N-acetylglucosamine deacetylase	1.36E-03	0.0038	1.5
lpxK	Z1261	Lipid A 4' kinase	8.77E-03	0.0093	1.8
lpxP	Z3643	Palmitoleoyl-acyl carrier protein (ACP)-dependent acyltransferase	7.19E-03	0.0084	1.5
lrp	Z1234	DNA-binding transcriptional dual regulator, leucine-binding	7.64E-03	0.0086	1.8

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
lsrA	Z2192	Fused AI2 transporter subunits of ABC superfamily: ATP-binding components	2.18E-02	0.0166	1.3
lsrB	Z2189	AI2 transporter	6.56E-04	0.0026	1.6
lsrC	Z2191	AI2 transporter	3.00E-03	0.0054	1.6
lsrD	Z2190	AI2 transporter	3.58E-04	0.0019	1.7
lsrK	Z2194	Autoinducer-2 (AI-2) kinase	4.68E-04	0.0022	2.1
ltaE	Z1104	L-allo-threonine aldolase, PLP-dependent	4.69E-03	0.0068	1.7
macA	Z1115	Macrolide-specific efflux protein	2.05E-04	0.0015	1.6
macB	Z1116	Putative ATP-binding component of a transport system	1.42E-03	0.0039	1.6
malI	Z2625	DNA-binding transcriptional repressor	4.45E-03	0.0066	1.3
manA	Z2616	Mannose-6-phosphate isomerase	1.68E-03	0.0042	1.5
marA	Z2170	DNA-binding transcriptional dual activator of multiple antibiotic resistance	1.83E-03	0.0043	1.4
marB	Z2169	Predicted protein	2.84E-03	0.0052	1.5
marC	Z2172	Predicted transporter	1.95E-03	0.0044	1.4
marR	Z2171	DNA-binding transcriptional repressor of multiple antibiotic resistance	6.01E-04	0.0024	1.3
mdtK	Z2690	Multidrug efflux system transporter	4.17E-03	0.0064	1.3
metB	Z5494	Cystathionine gamma-synthase, PLP-dependent	1.98E-02	0.0155	1.3
metF	Z5496	5,10-methylenetetrahydrofolate reductase	1.81E-03	0.0043	1.4
metJ	Z5493	DNA-binding transcriptional repressor, S-adenosylmethionine-binding	3.27E-02	0.0218	1.4
minC	Z1938	Cell division inhibitor	3.17E-04	0.0018	1.7
minD	Z1937	Membrane ATPase of the MinC-MinD-MinE system	1.17E-03	0.0035	1.7
minE	Z1936	Cell division topological specificity factor	1.74E-03	0.0043	1.5
modD	Z1962	Putative molybdenum transport protein	5.10E-04	0.0023	1.6
mqo	Z3468	Malate dehydrogenase, FAD/NAD(P)-binding domain	8.98E-03	0.0095	1.5
mraW	Z0092	S-adenosyl-dependent methyltransferase activity on membrane-located substrates	3.75E-03	0.0060	1.5
msbA	Z1260	Lipid transporter subunit of ABC superfamily; membrane component	1.31E-02	0.0119	2.0
mukB	Z1271	Predicted bifunctional chromosome partitioning protein and nucleotide hydrolase	1.35E-03	0.0038	1.6
mukE	Z1270	Protein involved in chromosome partitioning	2.73E-02	0.0192	1.5

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
mukF	Z1269	Involved in chromosome partioning, Ca2+ binding protein	8.62E-03	0.0092	1.6
murA	Z4552	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	4.79E-02	0.0285	1.2
murC	Z0101	UDP-N-acetylmuramate:L-alanine ligase	1.06E-03	0.0033	1.6
murD	Z0098	UDP-N-acetylmuramoyl-L-alanine:D-glutamate ligase	1.32E-03	0.0038	1.6
murE	Z0095	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate:meso-diaminopimelate ligase	7.37E-03	0.0085	1.4
murG	Z0100	UDP-N-acetylglucosamine:N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	3.60E-03	0.0059	1.7
mutT	Z0109	Nucleoside triphosphate pyrophosphohydrolase, marked preference for dGTP	8.88E-03	0.0094	1.3
nadB	Z3856	Quinolinate synthase, L-aspartate oxidase (B protein) subunit	2.04E-02	0.0158	1.5
napA	Z3463	Nitrate reductase, periplasmic, large subunit	3.02E-02	0.0206	1.3
napB	Z3460	Nitrate reductase, small, cytochrome C550 subunit, periplasmic	6.64E-03	0.0081	1.3
napC	Z3459	Nitrate reductase, cytochrome c-type,periplasmic	1.49E-02	0.0129	1.4
napD	Z3464	Assembly protein for periplasmic nitrate reductase	3.14E-02	0.0211	1.3
napF	Z3465	Ferredoxin-type protein, predicted role in electron transfer to periplasmic nitrate reductase (NapA)	1.88E-04	0.0014	1.5
napG	Z3462	Ferredoxin-type protein essential for electron transfer from ubiquinol to periplasmic nitrate reductase (NapAB)	1.03E-02	0.0102	1.3
napH	Z3461	Ferredoxin-type protein essential for electron transfer from ubiquinol to periplasmic nitrate reductase (NapAB)	1.30E-02	0.0119	1.3
nemA	Z2668	N-ethylmaleimide reductase, FMN-linked	1.57E-03	0.0041	1.6
nfsA	Z1078	Nitroreductase A, NADPH-dependent, FMN-dependent	3.08E-04	0.0018	1.8
nhaB	Z1949	Sodium:proton antiporter	1.60E-02	0.0135	1.6
nikD	Z4871	Nickel transporter subunit	2.01E-05	0.0004	2.0
nleA	Z6024	Non-LEE-encoded type III secreted effector	9.48E-03	0.0098	1.4
*nleG*2-1	Z6025	Predicted type III secreted effector (pseudogene)	2.94E-02	0.0202	1.3
*nleG*2-3	Z2149	Predicted non-LEE-encoded type III secreted effector	2.72E-02	0.0192	1.3
*nleG*2-4	Z2075	Predicted type III secreted effector (pseudogene)	1.16E-02	0.0111	1.5

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
nleG5-2	Z2151	Predicted non-LEE-encoded type III secreted effector	4.26E-02	0.0263	1.3
nlpI	Z4524	Lipoprotein	5.93E-05	0.0007	1.9
nrdA	Z3489	Ribonucleoside diphosphate reductase 1, alpha subunit	8.64E-03	0.0093	1.5
nth	Z2644	DNA glycosylase and apyrimidinic (AP) lyase (endonuclease III)	1.15E-03	0.0035	1.6
nudC	Z5571	NADH pyrophosphatase	2.59E-02	0.0185	1.3
nusA	Z4530	Transcription termination/antitermination L factor	3.51E-04	0.0019	1.4
obgE	Z4545	GTPase involved in cell partioning and DNA repair	2.12E-05	0.0004	2.3
ompC	Z3473	Outer membrane porin protein C	2.02E-03	0.0045	1.5
ompF	Z1276	Outer membrane porin	2.67E-03	0.0051	1.6
orf29	Z5102	LEE-encoded predicted type III secretion system factor	4.98E-03	0.0070	1.3
pdxA	Z0061	4-hydroxy-L-threonine phosphate dehydrogenase, NAD-dependent	6.01E-03	0.0076	1.3
pdxH	Z2652	Pyridoxine 5'-phosphate oxidase	9.70E-04	0.0032	1.3
pdxY	Z2648	Pyridoxal kinase 2/pyridoxine kinase	6.63E-03	0.0081	1.4
pepN	Z1280	Aminopeptidase N	1.62E-02	0.0135	1.4
pflC	Z5508	Pyruvate formate lyase II activase	3.14E-03	0.0055	1.3
pgaA	Z1526	Putative outer membrane poly-beta-1,6-N-acetyl-D-glucosamine (PGA) translocation/docking protein	1.30E-02	0.0119	1.5
pgaD	Z1523	Predicted inner membrane protein	1.05E-03	0.0033	1.8
phoA	Z0479	Bacterial alkaline phosphatase	4.06E-04	0.0021	1.3
pncB	Z1279	Nicotinate phosphoribosyltransferase	3.78E-03	0.0060	1.4
pntB	Z2597	Pyridine nucleotide transhydrogenase, beta subunit	2.68E-03	0.0051	1.3
polB	Z0068	DNA polymerase II	8.75E-03	0.0093	1.5
potF	Z1081	Putrescine transporter subunit: periplasmic-binding component of ABC superfamily	9.22E-03	0.0096	1.6
potH	Z1083	Putrescine transporter subunit: membrane component of ABC superfamily	4.36E-03	0.0065	1.7
potI	Z1084	Putrescine transporter subunit: membrane component of ABC superfamily	7.61E-03	0.0086	1.6
poxB	Z1105	Pyruvate dehydrogenase (pyruvate oxidase), thiamin-dependent, FAD-binding	1.40E-03	0.0038	1.8

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
proP	Z5713	Proline/glycine betaine transporter	1.16E-03	0.0035	1.5
proW	Z3980	Glycine betaine transporter subunit; membrane component of ABC superfamily	1.37E-04	0.0012	1.4
prrA	Z1961	Putative outer membrane receptor, probably *tonB* dependent	5.12E-04	0.0023	1.7
psiF	Z0480	Conserved protein	1.08E-03	0.0034	1.4
pspA	Z2482	Regulatory protein for phage-shock-protein operon	9.26E-08	0.0000	3.2
pspC	Z2479	*psp* operon regulatory protein	2.19E-06	0.0001	1.5
pstB	Z5216	Phosphate transporter subunit; ATP-binding component of ABC superfamily	7.48E-05	0.0008	1.5
pstS	Z5219	Phosphate transporter subunit; periplasmic-binding component of ABC superfamily	2.94E-10	0.0000	5.7
purR	Z2681	DNA-binding transcriptional repressor, hypoxanthine-binding	4.69E-04	0.0022	1.7
pyrE	Z5066	Orotate phosphoribosyltransferase	6.94E-04	0.0027	1.3
raiA	Z3890	Cold shock protein associated with 30S ribosomal subunit	3.65E-03	0.0059	1.3
rarD	Z5340	Predicted chloramphenical resistance permease	1.63E-02	0.0136	1.3
rcsB	Z3476	DNA-binding response regulator in two-component regulatory system with RcsC	1.62E-03	0.0042	1.5
rcsC	Z3477	Hybrid sensory kinase in two-component regulatory system with RcsB and YojN	2.81E-03	0.0052	1.8
rcsD	Z3475	Phosphotransfer intermediate protein in two-component regulatory system with RcsBC	4.42E-05	0.0006	1.7
recA	Z4002	DNA strand exchange and recombination protein with protease and nuclease activity	2.27E-06	0.0001	6.4
recG	Z5078	ATP-dependent DNA helicase	2.68E-02	0.0190	1.3
recO	Z3846	Gap repair protein	6.70E-03	0.0081	1.4
recX	Z4001	Inhibitor of RecA	6.37E-03	0.0079	1.4
ribB	Z4399	3,4-dihydroxy-2-butanone-4-phosphate synthase	6.94E-07	0.0001	2.7
ribC	Z2688	Riboflavin synthase, alpha subunit	6.04E-04	0.0025	1.7
rimK	Z1079	Ribosomal protein S6 modification protein	1.36E-02	0.0122	1.6
rluA	Z0066	Ribosomal large subunit pseudouridine synthase A	1.20E-02	0.0112	1.6
rluD	Z3888	23S rRNA pseudouridine synthase	1.15E-03	0.0035	1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
rluF	Z5620	23S rRNA pseudouridine synthase	3.75E-02	0.0240	1.4
rnk	Z0754	regulator of nucleoside diphosphate kinase	8.27E-04	0.0030	2.0
rnt	Z2671	Ribonuclease T (RNase T)	7.31E-04	0.0027	1.6
rplB	Z4688	50S ribosomal subunit protein L2	1.14E-02	0.0110	1.5
rplC	Z4691	50S ribosomal subunit protein L3	2.38E-04	0.0015	2.1
rplD	Z4690	50S ribosomal subunit protein L4	7.86E-06	0.0002	2.6
rplE	Z4678	50S ribosomal subunit protein L5	4.00E-02	0.0252	2.2
rplI	Z5812	50S ribosomal subunit protein L9	1.10E-02	0.0107	1.4
rplO	Z4671	50S ribosomal subunit protein L15	9.69E-04	0.0032	1.9
rplP	Z4684	50S ribosomal subunit protein L16	6.43E-04	0.0025	2.1
rplV	Z4686	50S ribosomal subunit protein L22	7.92E-05	0.0008	1.8
rplW	Z4689	50S ribosomal subunit protein L23	1.56E-02	0.0132	1.3
rplX	Z4679	50S ribosomal subunit protein L24	2.96E-02	0.0203	1.6
rpmA	Z4547	50S ribosomal subunit protein L27	1.78E-02	0.0144	1.3
rpmC	Z4683	50S ribosomal subunit protein L29	1.42E-02	0.0124	2.3
rpoA	Z4665	RNA polymerase, alpha subunit	1.81E-03	0.0043	1.9
rpoC	Z5561	RNA polymerase, beta prime subunit	2.79E-02	0.0195	1.4
rpsA	Z1257	30S ribosomal subunit protein S1	4.68E-04	0.0022	2.3
rpsC	Z4685	30S ribosomal subunit protein S3	8.63E-05	0.0009	2.1
rpsD	Z4666	30S ribosomal subunit protein S4	2.29E-04	0.0015	2.4
rpsE	Z4673	30S ribosomal subunit protein S5	5.98E-05	0.0007	2.4
rpsG	Z4699	30S ribosomal subunit protein S7	2.24E-05	0.0004	1.9
rpsH	Z4676	30S ribosomal subunit protein S8	1.87E-02	0.0149	2.2
rpsI	Z4588	30S ribosomal subunit protein S9	2.51E-05	0.0004	3.3
rpsJ	Z4692	30S ribosomal subunit protein S10	8.79E-04	0.0031	2.0
rpsK	Z4667	30S ribosomal subunit protein S11	1.96E-06	0.0001	2.1
rpsS	Z4687	30S ribosomal subunit protein S19	1.86E-04	0.0014	2.4

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
rraB	Z5867	Ribonuclease E inhibitor protein	8.81E-03	0.0093	1.3
rsd	Z5570	Regulator of sigma D	1.08E-02	0.0106	1.4
rsxA	Z2633	Inner membrane subunit of SoxR-reducing complex	3.27E-03	0.0056	1.6
rsxB	Z2634	Inner membrane iron-sulfur protein in SoxR-reducing complex	5.26E-03	0.0072	1.3
rsxC	Z2636	Inner membrane iron-sulfur protein in SoxR-reducing complex	3.28E-02	0.0218	1.4
rsxE	Z2642	Electron transport complex protein	1.40E-03	0.0038	1.5
rsxG	Z2640	Electron transport complex protein	8.73E-03	0.0093	1.3
rumB	Z1086	23S rRNA m(5)U747 methyltransferase	1.95E-03	0.0044	1.6
sbcC	Z0495	Exonuclease, dsDNA, ATP-dependent	5.25E-03	0.0072	1.3
sbcD	Z0496	Exonuclease, dsDNA, ATP-dependent	4.31E-02	0.0265	1.3
sbp	Z5462	Sulfate transporter subunit; periplasmic-binding component of ABC superfamily	1.96E-03	0.0044	1.3
sdhC	Z0875	Succinate dehydrogenase, membrane subunit, binds cytochrome b556	2.70E-03	0.0051	1.4
secA	Z0108	Preprotein translocase subunit, ATPase	7.69E-04	0.0028	1.6
secG	Z4537	Preprotein translocase membrane subunit	8.61E-03	0.0092	1.7
secM	Z0107	Regulator of *secA* translation	5.75E-05	0.0007	2.0
secY	Z4670	Preprotein translocase membrane subunit	4.76E-04	0.0022	2.1
serC	Z1253	3-phosphoserine/phosphohydroxythreonine aminotransferase	5.43E-04	0.0023	1.6
serS	Z1239	Seryl-tRNA synthetase, also charges selenocysteinyl-tRNA with serine	9.68E-03	0.0098	1.4
sgrR	Z0079	DNA-binding transcriptional regulator	8.81E-03	0.0093	1.6
slyA	Z2657	DNA-binding transcriptional activator	4.40E-03	0.0065	1.5
smtA	Z1268	Predicted S-adenosyl-L-methionine-dependent methyltransferase	4.07E-03	0.0063	1.7
sodC	Z2661	Superoxide dismutase, Cu, Zn	6.71E-03	0.0081	1.4
spr	Z3434	Predicted peptidase, outer membrane lipoprotein	2.84E-04	0.0017	1.5
ssuA	Z1284	Alkanesulfonate transporter subunit	2.49E-02	0.0180	2.0
ssuB	Z1281	Alkanesulfonate transporter subunit	2.94E-03	0.0053	1.8
ssuC	Z1282	Alkanesulfonate transporter subunit	1.28E-03	0.0037	1.8
ssuD	Z1283	Alkanesulfonate monooxygenase, FMNH(2)-dependent	1.39E-02	0.0123	1.7

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ssuE	Z1285	NAD(P)H-dependent FMN reductase	4.60E-02	0.0278	2.0
stx1A	Z3344	Shiga-like toxin 1 subunit A	1.36E-03	0.0038	1.3
stx1B	Z3343	Shiga-like toxin I subunit B	3.49E-03	0.0058	1.5
stx2A	Z1464	Shiga toxin II subunit A	7.90E-04	0.0029	1.3
sulA	Z1308	SOS cell division inhibitor	8.57E-06	0.0003	2.2
surA	Z0062	Peptidyl-prolyl cis-trans isomerase (PPIase)	6.92E-03	0.0082	1.4
tag	Z4974	3-methyl-adenine DNA glycosylase I, constitutive	1.88E-02	0.0150	1.3
tatD	Z5361	Mg-dependent DNase	1.07E-02	0.0104	1.5
tatE	Z0772	Component of Sec-independent translocase	7.64E-04	0.0028	1.8
tauD	Z0467	Taurine dioxygenase, 2-oxoglutarate-dependent	1.86E-03	0.0043	1.3
tbpA	Z0077	Periplasmic-binding component of thiamin ABC transporter	1.74E-02	0.0142	1.6
thiC	Z5569	Thiamin (pyrimidine moiety) biosynthesis protein	6.52E-03	0.0080	1.3
thiG	Z5565	Thiamin biosynthesis ThiGH complex subunit	6.13E-03	0.0077	1.3
thiL	Z0519	Thiamin-monophosphate kinase	2.24E-02	0.0169	1.4
thiP	Z0076	Membrane component of thiamin ABC transporter	2.44E-03	0.0049	1.5
thiQ	Z0075	ATP-binding component of thiamin ABC transporter	2.40E-03	0.0048	1.5
thiS	Z5566	Sulphur carrier protein	1.76E-03	0.0043	1.2
tolA	Z0907	Membrane anchored protein in TolA-TolQ-TolR complex	1.96E-02	0.0153	1.2
tolQ	Z0905	Membrane spanning protein in TolA-TolQ-TolR complex	7.41E-03	0.0085	1.3
torT	Z1411	Periplasmic sensory protein associated with the TorRS two-component regulatory system	2.00E-04	0.0014	1.3
treA	Z1968	Periplasmic trehalase (pseudogene)	2.08E-03	0.0045	1.6
trmD	Z3901	tRNA (guanine-1-)-methyltransferase	4.99E-03	0.0070	1.3
*trpE*_1	Z2546	Component I of anthranilate synthase	3.71E-02	0.0238	1.3
*trpE*_2	Z2547	Anthranilate synthase component I (EC 4.1.3.27)	1.02E-04	0.0010	1.4
trpS	Z4737	Tryptophan tRNA ligase	1.85E-02	0.0148	1.2

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
truB	Z4527	tRNA U55 pseudouridine synthase	6.91E-05	0.0008	2.0
trxB	Z1232	Thioredoxin reductase, FAD/NAD(P)-binding	2.56E-02	0.0184	1.5
ttk	Z5065	Putative transcriptional regulator	2.67E-03	0.0051	1.4
tufA	Z4697	Protein chain elongation factor EF-Tu (duplicate of *tufB*)	1.66E-04	0.0013	1.4
tyrS	Z2650	Tyrosyl-tRNA synthetase	4.29E-03	0.0065	1.4
ugpB	Z4822	Glycerol-3-phosphate transporter subunit	5.48E-03	0.0073	1.3
ulaA	Z5802	L-ascorbate-specific enzyme IIC component of PTS	1.02E-02	0.0102	1.5
ulaC	Z5804	L-ascorbate-specific enzyme IIA component of PTS	2.50E-02	0.0181	1.3
ulaE	Z5806	L-xylulose 5-phosphate 3-epimerase	1.08E-02	0.0105	1.4
ulaG	Z5801	Predicted L-ascorbate 6-phosphate lactonase	1.27E-02	0.0117	1.2
umuC	Z1947	DNA polymerase V, subunit C	7.34E-04	0.0027	2.0
umuD	Z1946	DNA polymerase V, subunit D	9.57E-03	0.0098	1.8
ung	Z3864	Uracil-DNA-glycosylase	1.49E-02	0.0129	1.5
uspG	Z0751	Universal stress protein UP12	1.52E-04	0.0012	2.2
uxaB	Z2184	Altronate oxidoreductase, NAD-dependent	7.06E-03	0.0083	1.3
wza	Z3227	Lipoprotein required for capsular polysaccharide translocation through the outer membrane	1.32E-02	0.0119	1.3
wzb	Z3226	Protein-tyrosine phosphatase	1.87E-02	0.0149	1.3
wzc	Z3224	Protein-tyrosine kinase	9.84E-05	0.0010	1.4
wzy	Z3203	O antigen polymerase	1.65E-02	0.0137	1.3
xdhD	Z4220	Fused predicted xanthine/hypoxanthine oxidase: molybdopterin-binding subunit	5.65E-03	0.0074	1.4
xylA	Z4990	D-xylose isomerase	4.33E-03	0.0065	1.4
xylB	Z4989	Xylulokinase	3.35E-02	0.0222	1.3
xylF	Z4991	D-xylose transporter subunit	1.54E-02	0.0131	1.3
xylG	Z4992	Fused D-xylose transporter subunits of ABC superfamily: ATP-binding	3.67E-02	0.0236	1.3
xylR	Z4994	DNA-binding transcriptional activator, xylose-binding	1.16E-02	0.0110	1.3
yabI	Z0074	Conserved inner membrane protein	2.86E-03	0.0052	1.5

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
yaiV	Z0470	Predicted DNA-binding transcriptional regulator	4.58E-02	0.0276	1.3
ybaE	Z0552	Predicted transporter subunit: periplasmic-binding component of ABC superfamily	4.84E-02	0.0288	1.3
ybbA	Z0648	Predicted transporter subunit: ATP-binding component of ABC superfamily	8.88E-05	0.0009	1.3
ybdA	Z0733	Putative transport	1.74E-03	0.0043	1.8
ybdB	Z0739	Conserved protein	2.72E-04	0.0017	2.0
ybdH	Z0742	Predicted oxidoreductase	2.90E-02	0.0200	1.7
ybdJ	Z0719	Predicted inner membrane protein	8.02E-04	0.0029	1.7
ybdK	Z0720	Gamma-glutamyl:cysteine ligase	8.73E-04	0.0030	2.0
ybdL	Z0743	Methionine aminotransferase, PLP-dependent	3.38E-06	0.0002	2.2
ybdM	Z0744	Conserved protein	9.40E-04	0.0031	1.8
ybdN	Z0746	Conserved protein	2.90E-04	0.0017	2.1
ybdO	Z0747	Predicted DNA-binding transcriptional regulator	2.00E-03	0.0045	1.8
ybdR	Z0752	Predicted oxidoreductase, Zn-dependent and NAD(P)-binding	1.90E-03	0.0044	2.0
ybeF	Z0774	Predicted DNA-binding transcriptional regulator	3.40E-03	0.0057	1.7
ybeM	Z0771	Carbon-nitrogen hydrolase	2.08E-03	0.0045	1.9
ybjC	Z1077	Predicted inner membrane protein	1.72E-03	0.0043	1.8
ybjD	Z1110	Conserved protein with nucleoside triphosphate hydrolase domain	9.32E-05	0.0009	2.0
ybjE	Z1108	Predicted transporter	5.96E-03	0.0076	1.6
ybjL	Z1074	Predicted transporter	2.80E-03	0.0052	1.4
ybjM	Z1075	Predicted inner membrane protein	2.09E-04	0.0015	1.6
ybjO	Z1085	Predicted inner membrane protein	1.26E-02	0.0117	1.6
ybjP	Z1095	Predicted lipoprotein	1.87E-03	0.0043	1.7
ybjQ	Z1099	Conserved protein	7.00E-03	0.0083	1.6
ybjR	Z1100	Predicted amidase and lipoprotein	4.45E-03	0.0066	1.6
ybjS	Z1102	Predicted NAD(P)H-binding oxidoreductase with NAD(P)-binding Rossmann-fold domain	7.38E-03	0.0085	1.6

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ybjT	Z1103	Conserved protein with NAD(P)-binding Rossmann-fold domain	4.00E-03	0.0062	1.8
ybjX	Z1112	Conserved protein	1.08E-03	0.0034	1.9
ycaC	Z1243	Predicted hydrolase	3.49E-02	0.0228	1.4
ycaD	Z1244	Predicted transporter	3.63E-02	0.0234	2.0
ycaJ	Z1238	Recombination protein	1.33E-02	0.0120	1.6
ycaL	Z1255	Predicted peptidase with chaperone function	3.61E-02	0.0233	2.3
ycaL	Z1255	Predicted peptidase with chaperone function	3.24E-03	0.0055	2.3
ycaM	Z1245	Predicted transporter	2.24E-03	0.0047	1.5
ycaO	Z1251	Conserved protein	7.22E-04	0.0027	1.8
ycaP	Z1252	Conserved inner membrane protein	3.94E-04	0.0020	1.8
ycaQ	Z1262	Conserved protein	4.27E-03	0.0065	1.5
ycaR	Z1263	Conserved protein	4.66E-03	0.0067	1.7
ycbB	Z1272	Predicted carboxypeptidase	6.75E-03	0.0081	1.3
ycbC	Z1267	Conserved inner membrane protein	2.14E-03	0.0046	1.6
ycbJ	Z1265	Conserved protein	5.22E-04	0.0023	1.5
ycbK	Z1273	Conserved protein	9.54E-04	0.0032	1.4
ycbL	Z1274	Predicted metal-binding enzyme	1.36E-03	0.0038	1.9
ycbQ	Z1286	Predicted fimbrial-like adhesin protein	1.71E-03	0.0043	1.8
ycbU	Z1291	Predicted fimbrial-like adhesin protein (pseudogene)	3.00E-02	0.0205	1.2
ycdC	Z1512	Putative *tet* operon regulator	1.47E-03	0.0040	1.5
ycdT	Z1527	Predicted diguanylate cyclase	2.28E-02	0.0171	1.4
yceD	Z1727	Conserved protein	5.75E-06	0.0002	1.6
ycgB	Z1951	Conserved protein	5.64E-04	0.0024	2.0
ycgJ	Z1939	Predicted protein	2.58E-04	0.0016	1.9
ycgK	Z1940	Predicted protein	2.11E-03	0.0046	1.8
ycgL	Z1941	Conserved protein	2.65E-04	0.0016	2.0
ycgM	Z1942	Predicted isomerase/hydrolase	9.46E-05	0.0009	1.9

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ycgN	Z1943	Conserved protein	1.07E-03	0.0034	1.9
*ycgR*_1	Z1959	N-terminal fragment of c-di-GMP binding protein (pseudogene)	1.73E-03	0.0043	1.9
ydeA	Z2173	Predicted arabinose transporter	8.63E-04	0.0030	1.4
ydeD	Z2168	Cysteine and O-acetyl-L-serine efflux system	4.97E-04	0.0023	1.3
ydeE	Z2166	Predicted transporter	1.33E-02	0.0120	1.3
ydeH	Z2163	Conserved protein	1.89E-04	0.0014	1.5
ydeI	Z2162	Conserved protein	5.95E-04	0.0024	1.5
ydeJ	Z2161	Conserved protein	2.99E-04	0.0018	1.5
ydeU	Z2196	C-terminal fragment of a predicted autotransporter (pseudogene)	5.42E-04	0.0023	1.7
ydfI	Z2155	Predicted mannonate dehydrogenase	2.82E-02	0.0196	1.3
ydfJ	Z2153	Predicted major facilitator superfamily (MFS) transporter (pseudogene)	4.71E-04	0.0022	1.4
ydgK	Z2632	Conserved inner membrane protein	4.28E-02	0.0264	1.3
ydhA	Z2653	Predicted lipoprotein	6.32E-03	0.0079	1.4
ydhB	Z2682	Predicted DNA-binding transcriptional regulator	1.24E-03	0.0036	1.4
ydhC	Z2685	Predicted transporter	3.40E-03	0.0057	1.5
ydhF	Z2664	Predicted oxidoreductase (pseudogene)	9.62E-03	0.0098	1.3
ydhI	Z2658	Predicted inner membrane protein	8.55E-04	0.0030	1.4
ydhK	Z2660	Conserved inner membrane protein	1.84E-03	0.0043	1.7
ydhO	Z2677	Predicted lipoprotein	1.79E-02	0.0145	1.4
ydhP	Z2679	Predicted transporter	8.57E-03	0.0092	1.3
ydhQ	Z2691	Conserved protein	1.17E-02	0.0111	1.3
ydhR	Z2694	Hypothetical protein	7.65E-04	0.0028	1.5
ydhS	Z2695	Conserved protein with FAD/NAD(P)-binding domain	1.35E-02	0.0121	1.4
ydhU	Z2697	Predicted cytochrome	2.37E-03	0.0048	1.5
ydhV	Z2701	Predicted oxidoreductase	2.50E-03	0.0049	1.5
ydhW	Z2700	Predicted protein	1.64E-03	0.0042	1.4
ydhY	Z2702	Predicted 4Fe-4S ferredoxin-type protein	2.55E-03	0.0050	1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ydhZ	Z2703	Predicted protein	2.62E-02	0.0187	1.3
ydjY	Z2783	Predicted protein	1.69E-02	0.0139	1.6
yebF	Z2899	Predicted protein	6.70E-05	0.0008	1.3
yegH	Z3229	Fused predicted membrane protein	1.51E-03	0.0040	1.5
yeiP	Z3430	Predicted translation elongation factor	7.71E-04	0.0028	1.3
yeiU	Z3433	Undecaprenyl pyrophosphate phosphatase	2.54E-03	0.0050	1.4
yfaA	Z3483	Predicted protein (pseudogene)	2.26E-03	0.0047	1.4
yfaD	Z3502	Conserved protein	1.57E-03	0.0041	1.4
yfaL	Z3487	Adhesin	2.42E-03	0.0049	1.4
*yfaQ*_1	Z3480	N-terminal fragment of a predicted protein (pseudogene)	1.14E-02	0.0110	1.6
yfaT	Z3482	Predicted protein	1.50E-03	0.0040	1.7
yfaV	Z3504	Predicted transporter	2.28E-03	0.0048	1.5
yfaW	Z3505	Predicted enolase	5.05E-03	0.0070	1.8
yfaX	Z3506	Predicted DNA-binding transcriptional regulator	1.16E-03	0.0035	1.7
yfaZ	Z3508	Predicted outer membrane porin protein	5.60E-04	0.0024	1.7
yfeC	Z3662	Predicted DNA-binding transcriptional regulator	2.69E-02	0.0190	1.3
yfeO	Z3655	Predicted ion channel protein	1.13E-02	0.0109	1.3
yfhL	Z3842	Predicted 4Fe-4S cluster-containing protein	1.87E-03	0.0043	1.4
yfiB	Z3899	Predicted outer membrane lipoprotein	3.86E-02	0.0245	1.3
yfiM	Z3871	Predicted protein	2.62E-02	0.0187	1.3
ygaW	Z3970	Predicted inner membrane protein	5.89E-07	0.0001	2.2
ygcM	Z4075	6-carboxy-5,6,7,8-tetrahydropterin synthase	1.87E-03	0.0043	1.3
ygdI	Z4126	Predicted protein	2.05E-05	0.0004	1.5
ygeG	Z4172	Predicted chaperone	7.10E-03	0.0083	1.4
ygeW	Z4209	Conserved protein	2.58E-02	0.0185	1.3
ygfQ	Z4223	Predicted transporter	3.12E-05	0.0005	1.5
ygfS	Z4224	Predicted oxidoreductase, 4Fe-4S ferredoxin-type subunit (pseudogene)	6.25E-03	0.0078	1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
ygfU	Z4226	Predicted transporter	8.53E-03	0.0092	1.3
ygiH	Z4412	Conserved inner membrane protein	9.43E-06	0.0003	1.3
ygjM	Z4435	Predicted DNA-binding transcriptional regulator	8.03E-05	0.0009	1.6
ygjN	Z4436	Conserved protein	1.01E-02	0.0101	1.5
ygjO	Z4437	Predicted methyltransferase small domain	2.52E-04	0.0016	1.5
ygjP	Z4438	Predicted metal dependent hydrolase	1.88E-02	0.0150	1.3
yhbC	Z4531	Conserved protein	3.19E-03	0.0055	1.3
yhbE	Z4546	Conserved inner membrane protein	4.10E-09	0.0000	5.4
yhbP	Z4514	Conserved protein	2.33E-04	0.0015	1.4
yhbV	Z4520	Predicted protease	3.15E-02	0.0211	1.3
yhbY	Z4542	Predicted RNA-binding protein	1.66E-04	0.0013	1.3
yhcB	Z4592	Conserved protein	1.50E-04	0.0012	1.4
yhdJ	Z4622	Predicted methyltransferase	4.39E-02	0.0269	1.3
yhdY	Z4631	Predicted amino-acid transporter subunit	1.19E-02	0.0112	1.3
yheS	Z4713	Fused predicted transporter subunits of ABC superfamily: ATP-binding components	1.94E-02	0.0153	1.3
yheU	Z4715	Conserved protein	2.31E-03	0.0048	1.4
yhgF	Z4762	Predicted transcriptional accessory protein	5.20E-03	0.0071	1.4
yhgN	Z4798	Predicted antibiotic transporter	1.56E-03	0.0041	1.3
yhhK	Z4831	Putative N-acetyltransferase	4.29E-04	0.0021	2.9
yhjX	Z4972	Predicted transporter	4.58E-04	0.0022	1.3
yiaA	Z4987	Conserved inner membrane protein	1.42E-02	0.0124	1.5
yiaB	Z4988	Conserved inner membrane protein	4.66E-03	0.0067	1.3
yiaD	Z4977	Predicted outer membrane lipoprotein	2.06E-02	0.0159	1.2
yiaH	Z4986	Conserved inner membrane protein	8.17E-03	0.0090	1.4
yibN	Z5038	Predicted rhodanese-related sulfurtransferase	2.27E-02	0.0170	1.3
yicG	Z5071	Conserved inner membrane protein	2.82E-05	0.0005	1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
yidC	Z5197	Cytoplasmic insertase into membrane protein, Sec system	7.09E-04	0.0027	1.4
yidE	Z5181	Predicted transporter	1.26E-02	0.0117	1.4
yieF	Z5208	Chromate reductase, Class I, flavoprotein	4.69E-02	0.0281	1.3
yiiX	Z5492	Predicted peptidoglycan peptidase	1.04E-02	0.0102	1.4
yjaG	Z5575	Conserved protein	2.67E-02	0.0189	1.3
yjaH	Z5577	Conserved protein	5.75E-03	0.0075	1.4
yjbO	Z5648	Phage shock protein G	8.80E-06	0.0003	3.1
yjeI	Z5749	Conserved protein	1.83E-02	0.0147	1.4
yjfC	Z5793	Predicted synthetase/amidase	4.35E-02	0.0267	1.3
yjfM	Z5792	Conserved protein	2.15E-02	0.0165	1.3
yjfN	Z5795	Predicted protein	1.05E-02	0.0104	1.3
yjfO	Z5796	Conserved protein	3.27E-04	0.0018	1.3
yjfP	Z5799	Palmitoyl-CoA and pNP-butyrate esterase	1.80E-03	0.0043	1.3
yjfY	Z5808	Predicted protein	1.86E-02	0.0148	1.3
ymgE	Z1960	Predicted inner membrane protein	5.43E-03	0.0073	1.5
*yneF*_2	Z2182	C-terminal fragment of a predicted diguanylate cyclase (pseudogene)	3.94E-05	0.0006	1.5
*yneK*_1	Z2176	N-terminal fragment of a predicted protein (pseudogene)	2.11E-04	0.0015	2.0
*yneK*_2	Z2175	C-terminal fragment of a predicted protein (pseudogene)	4.37E-03	0.0065	1.5
yneL	Z2199	Predicted transcriptional regulator	4.46E-03	0.0066	1.4
ynfG	Z2577	Oxidoreductase, Fe-S subunit	4.31E-02	0.0265	1.3
ynfL	Z2589	Predicted DNA-binding transcriptional regulator	1.96E-03	0.0044	1.3
ynhG	Z2706	Conserved protein	3.89E-04	0.0020	1.7
ynjA	Z2785	Conserved protein	1.50E-02	0.0129	1.5
ynjB	Z2786	Conserved protein	8.45E-03	0.0092	1.3
ynjC	Z2787	Fused transporter subunits of ABC superfamily: membrane components	8.18E-03	0.0090	1.4
yojI	Z3469	Fused predicted multidrug transport subunits of ABC superfamily: membrane component	9.31E-04	0.0031	1.7

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
yojL	Z3472	Predicted thiamine biosynthesis lipoprotein	3.35E-03	0.0057	1.5
ypdA	Z3645	Predicted sensory kinase in two-component system with YpdB	1.62E-02	0.0135	1.3
ypdB	Z3646	Predicted response regulator in two-component system with YpdA	2.33E-03	0.0048	1.4
yqaB	Z3991	Fructose-1-P and 6-phosphogluconate phosphatase	5.85E-03	0.0075	1.5
yqfA	Z4237	Predicted oxidoreductase, inner membrane subunit	1.70E-03	0.0043	1.4
yraQ	Z4510	Predicted permease	1.92E-02	0.0151	1.3
yrbC	Z4555	Predicted ABC-type organic solvent transporter	3.22E-03	0.0055	1.3
yrbD	Z4556	Predicted ABC-type organic solvent transporter	4.33E-05	0.0006	1.7
yrbL	Z4570	Predicted protein	4.34E-05	0.0006	1.4
ysaB	Z4985	Predicted protein	9.09E-03	0.0095	1.5
ytfH	Z5823	Predicted transcriptional regulator	1.48E-02	0.0128	1.3
ytfM	Z5831	Predicted outer membrane protein and surface antigen	6.40E-03	0.0079	1.4
ytfQ	Z5838	Predicted sugar transporter subunit: periplasmic-binding component of ABC superfamily	8.51E-03	0.0092	1.3
Z0271	Z0271	Predicted protein	6.17E-04	0.0025	1.3
Z0330	Z0330	Unknown protein encoded in prophage CP-933I	1.94E-02	0.0153	1.3
Z0635	Z0635	Putative membrane spanning export protein	2.37E-02	0.0174	1.3
Z0753	Z0753		1.79E-03	0.0043	2.1
Z1087	Z1087	Predicted protein	2.28E-03	0.0048	1.5
Z1088	Z1088	Predicted protein	6.17E-04	0.0025	1.7
Z1089	Z1089	Putative sulfatase	2.54E-03	0.0050	1.6
Z1098	Z1098	Predicted protein	2.10E-03	0.0046	1.8
Z1208	Z1208	Unknown in putative ISEc8	5.25E-04	0.0023	2.1
Z1217	Z1217	Putative DNA repair protein, RADC family	7.37E-03	0.0085	2.0
Z1371	Z1371	Putative tail assembly chaperone encoded by cryptic prophage CP-933M	8.90E-03	0.0094	1.3
Z1450	Z1450	Putative replication protein O of bacteriophage BP-933W	6.56E-04	0.0026	1.4
Z1452	Z1452	Unknown protein encoded by bacteriophage BP-933W	4.89E-03	0.0069	1.4

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
Z1453	Z1453	Unknown protein encoded by bacteriophage BP-933W	1.07E-05	0.0003	1.9
Z1454	Z1454	Putative DNA N-6-adenine-methyltransferase of bacteriophage BP-933W	1.66E-03	0.0042	1.5
Z1457	Z1457	Putative DNA-binding protein Roi of bacteriophage BP-933W	4.78E-02	0.0285	1.3
Z1459	Z1459	Antitermination protein Q of bacteriophage BP-933W	1.33E-04	0.0012	1.6
Z1480	Z1480	Unknown protein encoded by bacteriophage BP-933W	2.12E-03	0.0046	1.4
Z1495	Z1495	Unknown protein encoded by bacteriophage BP-933W	5.29E-03	0.0072	1.4
Z1506	Z1506	Predicted oxidoreductase, flavin:NADH component	6.76E-03	0.0081	1.4
Z1534	Z1534	Putative chaperone	4.40E-03	0.0065	1.3
Z1544	Z1544	Putative acyl-carrier protein synthase	3.78E-03	0.0060	1.4
Z1878	Z1878	Putative Bor protein of prophage CP-933X	3.39E-04	0.0019	1.4
Z1913	Z1913	Putative tail component of prophage CP-933X	1.75E-03	0.0043	1.7
Z1917	Z1917	Putative outer membrane protein of prophage CP-933X	4.76E-03	0.0068	1.9
Z1918	Z1918	Putative membrane protein of prophage CP-933X	1.36E-04	0.0012	2.0
Z1919	Z1919	Unknown protein encoded by prophage CP-933X	1.45E-03	0.0039	2.0
Z1921	Z1921	Unknown protein encoded by prophage CP-933X	4.20E-03	0.0064	2.1
Z1922	Z1922	Unknown protein encoded by prophage CP-933X	3.45E-03	0.0057	1.8
Z1923	Z1923	Unknown protein encoded by prophage CP-933X	4.61E-02	0.0278	1.3
Z1924	Z1924	Unknown protein encoded by prophage CP-933X	4.56E-04	0.0022	2.0
Z1925	Z1925	Unknown protein encoded by prophage CP-933X	1.49E-02	0.0129	1.6
Z1926	Z1926	Unknown protein encoded by prophage CP-933X	4.23E-03	0.0064	1.7
Z1931	Z1931	Outer membrane protease VII (outer membrane protein 3b)	2.00E-02	0.0156	1.5
Z1932	Z1932	Hypothetical protein	5.72E-04	0.0024	1.7
Z1957	Z1957	Transposase for IS629	1.56E-04	0.0012	1.9
Z1963	Z1963	Predicted protein	1.32E-04	0.0012	1.7
Z1964	Z1964	Putative iron compound ABC transporter, ATP-binding protein	2.16E-04	0.0015	2.0
Z1965	Z1965	Putative iron compound ABC transporter, permease protein	6.08E-04	0.0025	1.9
Z1966	Z1966	Predicted protein	1.55E-03	0.0041	2.0

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
Z1967	Z1967	Predicted protein	9.86E-04	0.0032	1.9
Z2038	Z2038	Unknown protein encoded by prophage CP-933O	6.67E-03	0.0081	1.5
Z2040	Z2040	Unknown protein encoded by prophage CP-933O	1.62E-02	0.0135	1.4
Z2045	Z2045	Similar to DicA, regulator of DicB encoded by prophage CP-933O	6.41E-04	0.0025	1.8
Z2047	Z2047	Unknown protein encoded by prophage CP-933O	1.89E-03	0.0044	1.5
Z2056	Z2056	Unknown protein encoded by prophage CP-933O	3.14E-03	0.0055	1.4
Z2065	Z2065	Similar to conserved hypothetical phage protein YjhS encoded within prophage CP-933O	5.39E-03	0.0072	1.3
Z2066	Z2066	Unknown protein encoded by prophage CP-933O	3.25E-03	0.0055	1.4
Z2083	Z2083	DinI family protein encoded within CP-933O	2.26E-03	0.0047	1.4
Z2084	Z2084	Putative integrase within CP-933O; partial	8.86E-03	0.0094	1.5
Z2091	Z2091	Unknown protein encoded within prophage CP-933O	8.16E-04	0.0029	1.4
Z2096	Z2096	Unknown protein encoded within prophage CP-933O	5.32E-03	0.0072	1.5
Z2098	Z2098	Unknown protein encoded within prophage CP-933O	1.51E-02	0.0130	1.3
Z2109	Z2109	Hypothetical phage protein similar to YjhS encoded within prophage CP-933O	1.91E-02	0.0151	1.3
Z2112	Z2112	Putative ClpP-like protease encoded within prophage CP-933O	4.24E-03	0.0065	1.5
Z2119	Z2119	Unknown protein encoded within prophage CP-933O	7.73E-03	0.0087	1.4
Z2121	Z2121	Unknown protein encoded within prophage CP-933O	2.07E-03	0.0045	1.3
Z2140	Z2140	Putative tail component of prophage CP-933O	2.87E-03	0.0052	1.4
Z2146	Z2146	Putative outer membrane protein Lom precursor of prophage CP-933O	6.42E-04	0.0025	1.4
Z2152	Z2152	Predicted protein	6.20E-03	0.0078	1.4
Z2165	Z2165	Rhodanese-like protein	2.60E-02	0.0186	1.3
Z2202	Z2202		1.63E-02	0.0135	1.5
Z2509	Z2509	Putative efflux pump	4.35E-02	0.0267	1.3
Z2542	Z2542	Predicted protein	1.21E-02	0.0113	1.4
Z2562	Z2562	Putative transposase (partial)	3.01E-02	0.0205	1.3
Z2631	Z2631	Predicted regulator	2.30E-02	0.0172	1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
Z2774	Z2774	Conserved protein	2.26E-02	0.0170	1.3
Z3230	Z3230	Predicted protein	1.89E-02	0.0150	1.3
Z3341	Z3341	Unknown protein encoded within prophage CP-933V	3.05E-03	0.0054	1.3
Z3345	Z3345	Putative antitermination protein Q for prophage CP-933V	3.19E-05	0.0005	1.3
Z3351	Z3351	Unknown protein encoded within prophage CP-933V	3.65E-02	0.0235	1.3
Z3481	Z3481	Predicted protein	3.99E-02	0.0251	1.4
Z3494	Z3494	Putative antibiotic efflux protein	3.15E-03	0.0055	1.5
Z3495	Z3495	Putative regulator	3.21E-03	0.0055	1.5
Z3496	Z3496	Predicted protein	3.86E-02	0.0245	1.3
Z3622	Z3622	Putative resolvase	5.15E-03	0.0071	1.3
Z3931	Z3931	Unknown protein encoded by prophage CP-933Y	1.37E-04	0.0012	1.8
Z4067	Z4067	Predicted protein	2.49E-03	0.0049	1.4
Z4069	Z4069	Hypothetical protein	6.02E-03	0.0076	1.4
Z4271	Z4271	Putative ATP-binding protein of ABC transport system	8.61E-03	0.0092	1.3
Z4383	Z4383	Putative iron compound permease protein of ABC transporter family	5.65E-04	0.0024	1.5
Z4385	Z4385	Putative ATP-binding protein of ABC transporter family	5.98E-04	0.0024	1.6
Z4777	Z4777	Hypothetical protein	7.47E-04	0.0028	1.3
Z4832	Z4832	Hypothetical protein	5.88E-05	0.0007	3.6
Z4833	Z4833	Hypothetical protein	2.52E-02	0.0182	1.3
Z4875	Z4875	Putative phosphotransferase system enzyme subunit	1.99E-04	0.0014	1.3
Z5093	Z5093	Unknown protein encoded within prophage CP-933L	2.35E-02	0.0174	1.4
Z5094	Z5094	N-terminal fragment of a predicted protein encoded within prophage CP-933L (pseudogene)	7.28E-03	0.0084	1.5
Z5095	Z5095	Internal fragment of a predicted protein encoded within prophage CP-933L (pseudogene)	2.79E-02	0.0195	1.3
Z5334	Z5334	Predicted protein	2.36E-02	0.0174	1.4
Z5335	Z5335	Predicted protein	1.37E-02	0.0122	1.3

Table 4.S1 (continued)

ID	Locus	Function/Product	*p value*	q value	FC
Z5339	Z5339	Predicted protein	4.25E-03	0.0065	1.3
Z5489	Z5489	Predicted protein	1.32E-02	0.0119	1.5
Z5490	Z5490	Predicted protein	2.57E-03	0.0050	1.4
Z5491	Z5491	Hypothetical protein in IS (partial)	3.49E-02	0.0228	1.4
Z5522	Z5522	Putative hippuricase	2.98E-02	0.0204	1.4
Z5614	Z5614	Putative sorbose PTS component	4.11E-03	0.0063	1.3
Z6069	Z6069	Putative DNA replication factor encoded within cryptic prophage CP-933P	2.73E-02	0.0192	1.5

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those who helped me over the duration of my graduate career. First and foremost, I would like to thank Dr. F. Chris Minion for his guidance, patience and support throughout this research and the writing of this thesis. Dr. Minion took a chance on an out-of-work undergraduate and molded him into the scientist and person I am today

I would also like to thank my committee members for their efforts, time, and contributions to this work: Dr. Nancy Cornick, Dr. Jeff Beetham, Dr. Larry Halverson and Dr. Dan Nordman.

I would like to thank all the former and current members of the Minion laboratory including but not limited to Melissa Madsen, Josh Pitzer, Janice Seibel, and Supraja Puttamreddy. Melissa Madsen designed and built the *E. coli* O157:H7 microarray along with teaching me microarray methodology. In essence, she built the foundation on which this research is built. For this I am grateful. Josh Pitzer was a researcher working in the Minion Laboratory untill the start of my graduate career. Josh taught me work ethic and humility. Our current Laboratory Manager, Janice Seibel, has been, for lack of a better phrase, a laboratory mother and has kept others and me on the right track. There were many Undergraduate Researchers, called Minions' Minion's Minions (in the Minion Laboratory), which were constantly working in the background. Without their contribution this work would have been painfully difficult. These are all extraordinary people.

There are many different people that have contributed directly to this work. Dr. Greg Phillips supplied pLacZY2 for the anaerobic studies and also *Acanthamoeba* for the *E. coli* O157:H7-protozoa studies. Dr. Thad Stanton and Sam Humphrey of the National Animal Disease Center along with Dr. Nancy Cornick taught and advised me on anaerobic techniques. Dr. Mike Wannemuehler allowed me to conduct the anaerobic work within his anaerobic chamber under the direction of Dr. Jennifer Wilson-Welder and Anne-Marie Overstreet. Jack Gallup gave invaluable advice and direction in qPCR.

I would like to thank my family for everything they have done and sacrificed for me over the years. Without their loving support I would not be where I am today.

Last, but certainly not least, I would like to acknowledge my wife Ashley who has been an endless source of love and support over these few years. She has given me so much and I only hope I can return the favor in years to come.