This is to certify that the
dissertation entitled

INTERPLAY BETWEEN ACID RESISTANCE AND VIRULENCE IN *ESCHERICHIA COLI* O157:H7

presented by

SIVAPRIYA KAILASAN VANAJA

has been accepted towards fulfillment
of the requirements for the

Ph.D degree in Comparative Medicine and Integrative Biology

Linda L. Mansfield
Major Professor's Signature

8/19/2009
Date

INTERPLAY BETWEEN ACID RESISTANCE AND VIRULENCE IN *ESCHERICHIA COLI* O157:H7

By

Sivapriya Kailasan Vanaja

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Comparative Medicine and Integrative Biology

2009

ABSTRACT

INTERPLAY BETWEEN ACID RESISTANCE AND VIRULENCE IN *ESCHERICHIA COLI* O157:H7

By

Sivapriya Kailasan Vanaja

Escherichia coli O157:H7 is a food borne pathogen of zoonotic origin that causes hemorrhagic colitis and hemolytic uremic syndrome in humans. Acid resistance (AR) is an essential characteristic of pathogenic *E. coli* O157:H7 because during its transit from reservoir host, cattle, to humans, O157 has to survive acidic environments of different food vehicles and the human stomach. *E. coli* O157:H7 can survive in extreme acidic conditions such as pH 1.5 – 2.5 of human stomach with the help of four AR mechanisms. Glutamate decarboxylase (GAD) system is the most efficient of the four AR mechanisms and is regulated by a central activator, GadE. It is possible that while integrating the laterally acquired genes into the chromosomal regulatory network of *E. coli* O157:H7, chromosomal regulators such as GadE has evolved into a global regulator with additional functions besides regulating GAD system. However, role of GadE on a genome wide scale remains poorly defined in *E. coli* O157:H7. The effect of AR induction on the expression of virulence factors of *E. coli* O157:H7 is not clear. When exposed to acidic conditions, *E. coli* O157:H7 may downregulate its virulence mechanisms to conserve energy. Moreover, it is possible that differential expression of AR and virulence related genes may be the reason for the variation in infectivity observed between different genotypes of *E. coli*

O157:H7. In this context, the specific aims of this project were: 1) to elucidate the role of GadE in acid resistance and virulence of *E. coli* O157:H7, 2) to determine the effect of acidic pH on the expression of the LEE pathogenicity island and to understand the role of GAD system regulators in this effect, and 3) to determine whether the differences in infectivity of clinical and bovine-biased genotypes of *E. coli* O157:H7 is due to the differential expression of virulence and stress fitness associated genes. To address these goals transcriptional profiling of wild type and mutants of GAD system regulators was conducted. Expression studies were also conducted to identify the genes differentially expressed between clinical and bovine-biased genotypes of *E. coli* O157:H7. Understanding the molecular mechanisms of AR and their effect on virulence and pathogenesis of *E. coli* O157:H7 could help in developing new methods for diagnosis, prevention and control of this pathogen. Moreover, identifying the differentially expressed genes between different genotypes of *E. coli* O157:H7 with varying infectivity would help us to identify new targets for vaccine development.

This work is dedicated to the fond memories of

Thomas S. Whittam, my mentor.

ACKNOWLEDGEMENTS

I wish to thank my mentor, Thomas S. Whittam, for being a great advisor and for his valuable advice and support, and for providing the resources and opportunities for me to acquire competence in a broad range of topics. Linda Mansfield, my acting advisor, has been a great support during the final stages of my PhD and I thank her for her guidance. I thank my committee members, John Linz, Robert Britton and the CMIB program director, Vilma Yuzbasian-Gurkan, for their support, suggestions, and guidance.

Many of my current and previous lab mates have provided suggestions, support, and technical assistance that were critical to the completion of this research. I thank Teresa Bergholz, Shannon Manning, Galeb Abu-Ali, James Riordan, Adam Nelson, Nicole Milton, Hans Steinsland, Lindsey Ouellette, Cody Springman, Jillian Tietjen, Mahesh Neupane and Scott Henderson.

I wish to thank my parents, Kailasan Krishnan and Vanaja Parameswaran, my sister, Jayapriya Anil, and my best friend, Mahrukh Banday, for their support and encouragement. My husband Vijay Rathinam has provided me with amazing support during my graduate career. I thank him for his love and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

PATHOGENIC *ESCHERICHIA COLI*

Escherichia coli is a genetically diverse group of Gram-negative facultative anaerobe that typically colonizes the large intestine and lower part of the small intestine of mammals (66). Commensal *E. coli* coexist with the human host without causing any disease whereas pathogenic forms of *E. coli* can cause a number of clinical illnesses with varying degrees of severity (8, 66). Acquisition of several virulence factors through horizontal gene transfer during evolution has allowed pathogenic *E. coli* to adapt to new predilection sites in hosts and to successfully cause a broad spectrum of diseases in them. Different combinations of these acquired virulence factors defines specific pathotypes of *E. coli*, which can cause characteristic clinical symptoms such as enteric disease, renal and urinary tract infections, and meningitis (66).

Clinically important enteric pathotypes of *E. coli* include enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), Shiga toxin (Stx)-producing *E. coli* (STEC), and enterohemorrhagic *E. coli* (EHEC). Other enteric pathotypes are enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC), and diffusely adherent *E. coli* (DAEC). ETEC is the leading cause of traveler's diarrhea in developing countries and it harbors at least one enterotoxin; heat-labile toxin (LT) or heat-stable toxin (ST). EPEC causes diarrhea in children younger than 2 years and the infection is characterized by the attaching and effacing (A/E) lesions in the intestinal mucosa. This lesion is mediated by a laterally acquired pathogenicity island, the locus of enterocyte effacement (LEE). *E. coli* strains harboring one or more variants of Stx are named STEC, and EHEC is a sub

population of STEC defined by the presence of Stx, LEE and a pO157 plasmid. EHEC infections are characterized by hemorrhagic colitis and hemolytic uremic syndrome (HUS) (66, 101, 109).

E. coli O157:H7

E. coli O157:H7 is the most prevalent EHEC serotype in the United States (153). It is a zoonotic food-borne pathogen and causes hemorrhagic colitis and HUS in humans similar to other EHEC. *E. coli* O157:H7 was isolated first in 1983 from stool cultures of patients with hemorrhagic colitis associated with ingestion of undercooked hamburgers and from sporadic cases of HUS (153). It has since become an emerging pathogen and currently, O157 is estimated to cause approximately 73,000 illnesses in the US yearly with an economic burden of 0.2 – 0.6 billion (120).

E. coli O157:H7 strains are distinguished from other serotypes by their inability to ferment sorbitol (SOR^-) and to produce β-D-glucuronidase (GUD^-) (101). A SOR^+ GUD^+ EPEC serotype, *E. coli* O55:H7, is considered as the evolutionary ancestor of *E. coli* O157:H7 (121). *E. coli* O157:H7 was evolved from O55:H7 in a step-wise manner through acquisition of Stx2 and the pO157 plasmid; followed by an antigenic shift, gain of Stx1, and loss of motility, SOR^+, and GUD^+ phenotypes (Fig. 1.1) (170). Whole genome sequence comparison of *E. coli* O157:H7 Sakai with benign laboratory strain *E. coli* K-12 revealed that the O157 genome contains 1.4-Mb of O157:H7-specific sequence most of which is

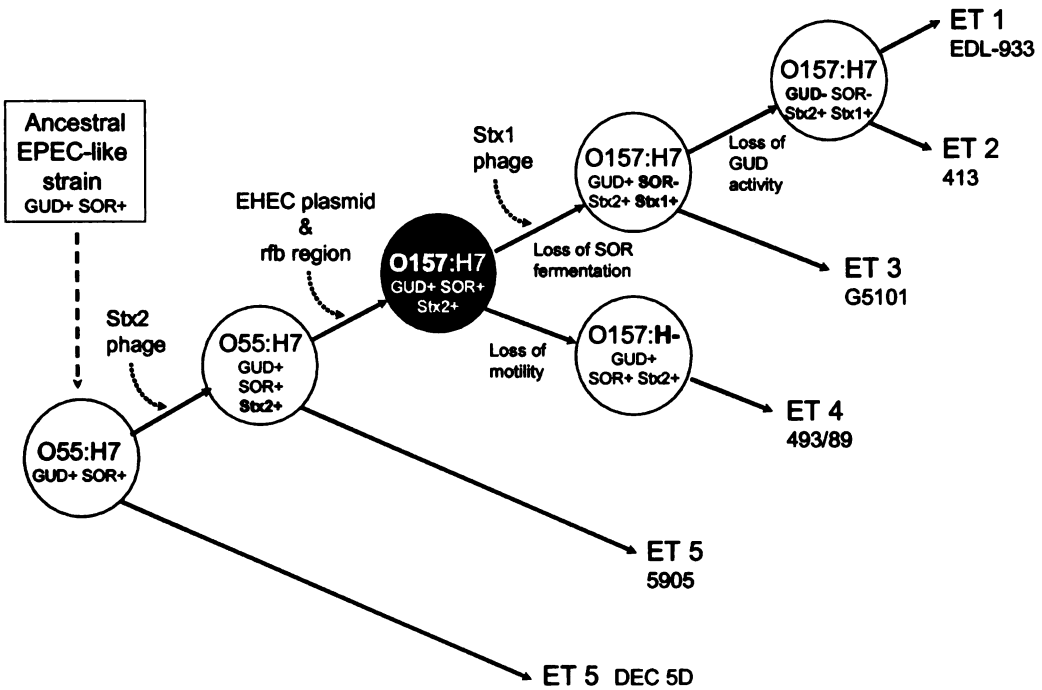


Figure 1.1. Step-wise evolution of *E. coli* O157:H7 from EPEC-like ancestor. Phenotypes of ancestors are shown (38).

contributed by foreign DNA elements that are acquired horizontally during evolution (46).

Clinical manifestations and treatment

The infective dose of *E. coli* O157:H7 is extremely low: 10 –100 cells are sufficient to cause a clinical infection (68, 147, 163). The average incubation period for *E. coli* O157:H7 infection is 3.7 days. The disease starts as non-bloody diarrhea, which lasts for 1 – 3 days followed by bloody diarrhea. Bloody diarrhea occurs in 90% of the cases due to severe hemorrhagic colitis. In O157 infections, most patients remain afebrile and the abdominal pain is more severe compared to other bacterial gastroenteritis (153). The most lethal complication of *E. coli* O157:H7 infection is HUS, which causes acute renal failure in 15% of the infected children younger than 10 years old. Onset of HUS occurs typically between 5 – 13 days of the infection and generally HUS starts as thrombocytopenia followed by hemolysis and azotaemia (153). HUS is a thrombotic disorder and the characteristic lesions include microvascular thrombi and swollen endothelial cells (57, 160).

It is believed that intravenous rehydration and maintenance fluids provide optimal protection against kidney damage and thus, constitute the most common management strategy in patients with bloody diarrhea. Expansion of parenteral volume has been associated with attenuation of renal injury (3). However, standard rehydration protocols are considered inadequate for O157 infections. Isotonic crystalloid is highly recommended for volume expansion and

maintenance in HUS. Antibiotics are not administered to patients with O157 infections as many studies have shown a strong correlation between antibiotic therapy and increased risk of HUS (128, 172).

Reservoir hosts of *E. coli* O157:H7

Cattle are the primary reservoir hosts of *E. coli* O157:H7 and O157 exists as part of normal intestinal microflora of cattle without causing any disease (25, 74). In addition, other food animals such as pigs, sheep, and goats (104) act as reservoir hosts of *E. coli* O157:H7 (106). *E. coli* O157:H7 colonizes the gastrointestinal tract of cattle, specifically the lymphoid follicle-dense mucosa of the recto-anal junction (40, 102). Earlier surveys in cattle indicated lower prevalence of *E. coli* O157 in feces possibly due to the poor sensitivity of isolation methods (49). Only 1.8% of fecal samples were found to contain *E. coli* O157 in one of the largest surveys. However, recent studies have indicated a markedly higher prevalence of O157 in cattle (51, 62). An investigation on the correlation of *E. coli* O157 prevalence in feces, hides and carcasses of beef cattle during processing demonstrated that of the 30 lots sampled, 87% had at least one O157 positive pre-evisceration sample (36). Currently, *E. coli* O157:H7 is considered ubiquitous in cattle farms and the shedding rate in cattle farms is estimated to be greater than 10%, sometimes approaching 100% (14, 45). Interestingly, there is a seasonal effect in the prevalence of *E. coli* O157:H7 in cattle with peak prevalence in summer and early fall (45). This corresponds with the peak in outbreaks involving ground beef in summer (120).

Transmission of *E. coli* O157:H7

A factor critical for the pathogenicity of *E. coli* O157:H7 is its ability to be transmitted from cattle to humans through a variety of food vehicles. In the earlier outbreaks O157 was typically transmitted through contaminated ground beef (115, 118). Later it was found to be transmitted through vehicles such as apple cider (13, 53), dry salami (116), apple juice (29), and raw milk (114). Most recently, green leafy vegetables such as spinach and lettuce are implicated in O157 outbreaks (52, 117). Interestingly, many of these vehicles pose multiple environmental stresses and *E. coli* O157:H7 appears to be successful in surviving under these conditions. Response of O157 to the acidity of apple juice (pH 3.5) has been investigated using a model apple juice (MAJ) system and it was found that O157 is induced for multiple stress response regulons including the RpoS, RpoH and CpxRA regulons in response to MAJ exposure (9).

VIRULENCE FACTORS AND PATHOGENESIS

Important virulent factors of *E. coli* O157:H7 include Stx, LEE and the pO157 plasmid. Each of these factors and their role in pathogenesis and regulation are described in detail below.

(i) Stx: Stx, also known as verotoxin (VT), is encoded by bacteriophages that are inserted into the O157 chromosome (66). There are two subgroups in the Stx family: Stx1 and Stx2 with five allelic variants of Stx2 (Stx2, Stx2c, Stx2d, Stx2e and Stx2f) (153). There is a 55% amino acid homology between Stx1 and Stx2. Stx is a typical AB_5 exotoxin with five identical B subunits forming a pentamer with a single A subunit. The A subunit of Stx has two peptide chains A1 and A2. A1 is a N-glycosidase enzyme and A2 connects A1 with the B-pentamer (66). The B-pentamer mediates the binding of Stx to the glycolipid globotriaosylceramide (Gb3) receptors on the host cell surface. This binding is followed by a receptor-mediated endocytosis and transfer through the Golgi apparatus and endoplasmic reticulum. During this transfer, a membrane bound protease, furin, nicks the A subunit from the holotoxin leaving the A1 fragment attached to A2 by a disulfide bond. This bond is later reduced and subsequently, the A1 enzyme is translocated into the cytoplasm where it removes an adenine residue from the 28s rRNA and inhibits the elongation step of protein synthesis. Inhibition of protein synthesis leads to death of cells that carry Gb3 receptors. There are cellular differences in the expression of Gb3 and renal glomerular endothelial, mesangial and tubular epithelial cells are characterized by higher

concentration of Gb3 receptors, which makes them highly susceptible to the damage by Stx. Direct toxicity by Stx along with induction of cytokines and chemokines cause destruction of renal endothelial cells and microvascular occlusion ultimately leading to HUS. Stx also causes intestinal epithelial cell death and damage in the colon leading to hemorrhagic colitis and bloody diarrhea by a similar mechanism (7).

Stx is encoded on a single operon comprising two genes corresponding to the A and B subunits. The 1:5 ratio of the A/B subunit synthesis for generating the AB_5 holotoxin is maintained by a stronger ribosomal binding efficiency of the B-subunit gene resulting in increased translation of B-subunits compared to the A-subunit (109). Stx genes are located on the late gene region of the Stx-producing lambdoid prophage down stream of the phage lambda Q antiterminator gene. Increased toxin synthesis occurs upon induction of the prophage due to an increase in copy number as the phage genome replicates and due to upregulation of stx expression from the phage late promoter P_R'. Stx1 has its own promoter in addition to the phage late promoter and this promoter can induce toxin gene transcription independent of phage induction. No phage independent promoter has been identified for Stx2. Genes encoding holin and endolysin are located down stream of the Stx genes and therefore, it is believed that coupling of toxin and lysis gene expression facilitates release of the toxin from the bacterial cell (149).

Several environmental factors that regulate Stx expression have been identified. There is a strong correlation between DNA damage and Stx phage

induction. Antibiotics such as norfloxacin and ciprofloxacin inhibit cellular DNA gyrase activity resulting in DNA damage and SOS response, which lead to a dramatic increase in Stx production from cells (48). Iron downregulates Stx1 expression as the ferrous iron binding protein, Fur, directly binds to the Fur box upstream of the phage independent Stx1 promoter and represses its transcription (20, 95).

(ii) LEE: LEE is a laterally acquired pathogenicity island that encodes a type 3 secretion system (T3SS) in EHEC and EPEC. The LEE consists of 41 genes that are transcribed as five polycistronic operons, LEE1 through LEE5. LEE1 encodes the key regulator of LEE, the Ler (LEE encoded regulator), which positively regulates the expression of LEE2 through LEE5. LEE also encodes additional regulators such as GrlA (global regulator of LEE activator) and GrlR (global regulator of LEE repressor). The structural components of T3SS are mainly encoded on the 5′ end of LEE whereas the outer membrane adhesin intimin (Eae) and the translocated intimin receptor (Tir) are encoded on the central part of LEE. The 3′ end of LEE encodes translocators, effectors and additional structural proteins of the T3SS. LEE-encoded T3SS is involved in translocating effector proteins encoded on the LEE and on different locations in the genome (158), into the host cell cytosol, which is important for the pathogenesis of O157 infection (41, 95).

The T3SS apparatus assembled from the products of LEE genes has the typical multicomponent organelle structure of gram-negative T3SS with outer and

inner membrane ring structures and a needle complex (41). This T3SS is responsible for the characteristic intimate attachment between *E. coli* O157:H7 and intestinal epithelial cells. Following initial adherence of *E. coli* O157:H7 to the host cell, the T3SS needle apparatus is inserted into the host cell plasma membrane and Tir is secreted into the cytoplasm, which acts as a receptor for intimin. Binding between intimin and Tir leads to an intimate attachment of bacteria to host cell. This induces rearrangement of actin cytoskeleton resulting in formation of pedestal-like structures and effacement of intestinal microvilli ultimately causing A/E lesions in the intestine, the hallmark of O157 infection.

Regulation of LEE expression is complex and involves multiple environmental factors and regulators (95). Growth in tissue culture medium such as Dulbecco's modified Eagle's medium (DMEM) at host body temperature (37°C), pH 7.0, and physiological osmolarity induces maximal LEE expression (69, 70). Other environmental factors that stimulate LEE expression include presence of iron and sodium bicarbonate. Presence of ammonium chloride and exclusion of calcium from the growth medium can inhibit LEE expression (56). As mentioned above, Ler is the key positive regulator of LEE and it acts by disrupting the silencing of LEE by the histone-like DNA binding protein, H-NS (95). Several regulatory systems such as RcsCDB phosphorelay system and the EHEC-specific GrvA protein regulate expression of LEE by controlling the *ler* expression in O157 (157). In addition, there is a quorum-sensing regulation of LEE, which involves an AI-3-signaling molecule that cross talks with epinephrine hormone and activates LEE expression (141, 142).

(iii) pO157 plasmid: pO157 is a laterally acquired 92 kb F-like plasmid present in all *E. coli* O157:H7 strains (19). pO157 contains 100 open reading frames (ORFs) and encodes a number of virulence factors such as enterohemolysin (EhxA), toxin B (ToxB) and a type 2 secretion system (T2SS) encoded by *etp* genes (19). Enterohemolysin is a repeats in toxin (RTX) family toxin. Although the exact role of enterohemolysin in the pathogenesis of O157 infections is not clear, it has been shown to cause endothelial injury and induce production of cytokine, IL-1β, which is a hallmark of HUS (152). *ehxA* is the structural gene of enterohemolysin, which is encoded on a single operon containing four genes *ehxCABD*. EhxC converts EhxA into an active form by the addition of a fatty acid group (58) and EhxB and D form a secretion machinery for EhxA (166). Another virulence factor encoded by pO157, ToxB, is a potential adhesin with sequence similarity to the *Chlostridium* toxin family. The *etp*-T2SS secretes StcE, a zinc metalloprotease that cleaves the C1 esterase inhibitor of the complement pathway, which may contribute to the tissue damage (76). StcE also has a mucinase activity and is involved in intimate adherence of O157 to host cells (44).

A positive interaction exists between LEE and many of the pO157 encoded virulence genes. GrlA, the LEE encode regulator, positively regulates the transcription of *ehxA*. Similarly, Ler upregulates the expression of *stcE* (44). Furthermore, it has been shown that mutation of the *toxB* gene results in reduced

expression and secretion of proteins encoded by LEE leading to a decrease in adherence to cultured epithelial cells (143).

ACID RESISTANCE (AR) OF *E. COLI* O157:H7

Gastric acidity (pH 1.5–2.5) is one of the first innate defense barriers encountered by enteric pathogens upon entry into the human body (112). Enteric pathogens have adapted several mechanisms to breach this defense barrier and establish infection. One common strategy is the 'assault tactic' as seen in *Vibrio cholerae* and *Salmonella* infections, where the infectious dose of the organism is so huge that some organisms will survive and enter the intestine. Another strategy shown typically by *Helicobacter pylori* is to mount a counterdefence mechanism such as the urease system to neutralize the extreme acidity (39, 135, 137). *E. coli* O157:H7 also has to overcome the adverse conditions of the stomach before successfully colonizing its niche, the large intestine. Similar to *H. pylori*, *E. coli* strains have been shown to be more AR than other enteric pathogens (83). This extreme AR is governed by four principal mechanisms in *E. coli*; the oxidative (OXI) system, the GAD system, the arginine decarboxylase (ARG) system and the lysine decarboxylase (LYS) system (39, 83). The OXI system is the least understood AR system in *E. coli*. It is induced upon entry into the stationary phase in a complex medium at pH 5.5. Once induced, the OXI system helps the bacteria to survive in pH 2.5 in minimal medium. Global regulatory protein CRP (cAMP receptor protein) and the stationary phase sigma factor, RpoS are essential for the functioning of the OXI system (23, 82). Presence of glucose in the medium inhibits CRP and hence, represses the OXI system (23).

Functioning of GAD, ARG and LYS systems are dependent on the availability of the amino acids glutamate, arginine and lysine, respectively. These systems typically consist of pairs of amino acid decarboxylases and antiporters. The ARG system comprises the arginine decarboxylase enzyme encoded by *adiA* and the arginine-agmantine antiporter encoded by *adiC* (59). Anaerobic conditions in a complex medium at low pH are essential for transcription of *adiA* and *adiC*. It is believed that a regulator, CysB, senses these factors and induces their transcription. An AraC-like regulator, AdiY, also activates the expression of *adiA* and *adiC*, however, this regulator is not essential for their expression (43, 145). Similarly, RpoS is involved in the functioning of the OXI system, but is not required for the transcription of *adiA* and *adiC* genes. The LYS system is a much less efficient system compared to the other three AR systems. It consists of lysine decarboxylases (*cadB*) and the lysine-cadaverine antiporter (*cadA*) and provides protection against mild acidity (59).

GAD system

The GAD system is the most efficient AR mechanism in *E. coli* and provides best protection against pH as low as 2.0 (75). The components of the GAD system include two glutamate decarboxylase isozymes, GadA and GadB, and a membrane-associated glutamate-gamma amino butyric acid (GABA) antiporter, GadC (23, 24, 75). GadA and GadB are pyridoxal 5'-phosphate (PLP)-dependent enzymes and have a hexameric structure with one PLP moiety per monomer (16). The nucleotide sequences of GadA and GadB are 98% similar

and they have highly similar peptide structure with differences only in 5 amino acid residues mainly at the N-terminal region (140). The two isozymes also share similar biochemical characteristics such as specific activity and isoelectric points (16). The two genes *gadA* and *gadB* map to distinct loci in the *E. coli* K-12 chromosome with *gadB* forming an operon with *gadC* at 33.8 min and *gadA* at 78.98 min as part of the acid fitness island. Either one of the decarboxylases isozymes is sufficient for GAD dependent AR at pH 2.5 whereas both are required for AR at pH 2.0. Presence of GadC is essential for the functioning of GAD system at both conditions (23).

The GAD system is induced upon entry into stationary phase of growth as well as when the cells are exposed to acidic conditions. Presence of glutamate in the environment is essential for the functioning of the GAD system. As the extracellular pH decreases, the intracellular pH of *E. coli* also goes down. At an intracellular pH of 4.2, GadA and GadB get activated and these enzymes catalyze the replacement of the carboxyl group of glutamate with a proton from the cytoplasm producing GABA and CO_2. This reaction reduces the intracellular H^+ ion concentration and stabilizes the pH homeostasis of the cell and maintains the pH at 4.2, which is tolerable to the *E. coli* physiology (39, 123). GABA is in turn expelled into the extracellular environment via GadC, in exchange for incoming glutamate (Fig. 1.2) (120).

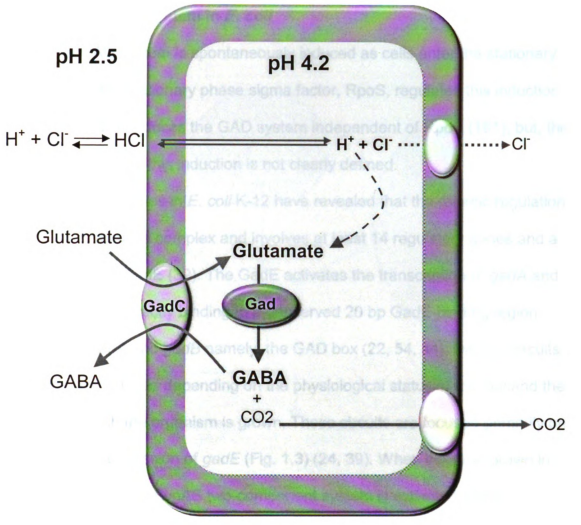


Figure 1.2. Mechanism of action of GAD system. GABA= gamma amino butyric acid. Adapted from (39).

Regulation of GAD system in *E. coli*

The GAD system is spontaneously induced as cells enter the stationary phase of growth. Stationary phase sigma factor, RpoS, regulates this induction (87). Low pH also induces the GAD system independent of RpoS (161), but, the exact mechanism of this induction is not clearly defined.

Previous studies in *E. coli* K-12 have revealed that the genetic regulation of the GAD system is complex and involves at least 14 regulatory genes and a central activator, GadE (39). The GadE activates the transcription of *gadA* and *gadBC* genes by directly binding to a conserved 20 bp GadE-binding region upstream of *gadA* and *gadB* namely, the GAD box (22, 54, 84). Distinct circuits of regulation are induced depending on the physiological status of the cell and the medium in which the organism is grown. These circuits are focused primarily on regulating the expression of *gadE* (Fig. 1.3) (24, 39). When cells are grown in minimal medium, the EvgAS two-component system is activated and it upregulates the expression of YdeO, which in turn triggers the expression of *gadE* (85, 89, 90). On the other hand, in a complex medium such as LB, during the stationary phase of growth, another circuit comprised of CRP, RpoS and two AraC-like regulators, GadX and GadW comes into play. RpoS induces the expression of GadX/W, which in turn indirectly activates the transcription of *gadA/B* via GadE (87, 159, 162). GadX and GadW also form a negative feed back loop by binding to the GAD box sequences and repressing the *gadA* and *gadBC* promoters. Because of this dual role, GadX/W can act as activators of

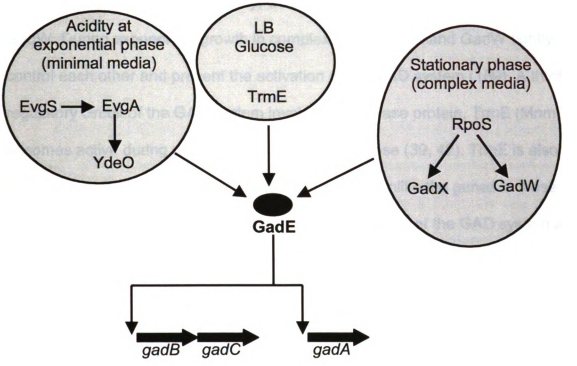


Figure 1.3. Regulatory circuits involved in the induction of GAD system. The growth phase and medium that activate each circuit is given within the circles. Modified from (39).

glutamate decarboxylase under some conditions and repressors under others (87). GadW forms another feed back loop by inhibiting RpoS, which is necessary for the activation of GadX. Similarly, GadX can repress the transcription of GadW. During exponential growth in complex media GadX and GadW tightly control each other and prevent the activation of the GAD system (159). A third regulatory circuit of the GAD system involving a GTPase protein, TrmE (MnmE), becomes active during growth in LB containing glucose (39, 42). TrmE is also required for efficient translation of GAD genes (42). Unlike the genes in these regulatory circuits, GadE is necessary for the functioning of the GAD system at all stages of growth in any medium (39, 84).

In addition to these regulatory circuits, other regulators such as RcsB and RNaseE are also involved in the regulation of GAD system. Basal level expression of RcsB, a component of RcsCDB signal transduction system, is necessary for GAD system functioning whereas, if over-expressed, RcsB represses GAD expression (24). RNaseE, an essential endoribonuclease involved in processing and degradation of RNAs, is also required for induction of the GAD system (150). Similarly, a functional topoisomerase I is essential for full induction of the GAD system (144). PhoP, the response regulator of the PhoQ/PhoP two-component system, also positively regulates the GAD system by directly binding to upstream sequences of GadE and GadW (176). Besides above-mentioned protein regulators, a small RNA, GadY, is also involved in induction of the GAD system. GadY is encoded by the AFI, on the opposite strand between *gadX* and *gadW*. There are three forms of GadY, a 105

nucleotide long form and two processed forms of 90 and 59 nucleotides. RpoS controls GadY expression and it is induced at stationary phase of growth. GadY acts by conferring increased stability to *gadX* mRNA by forming base pairs with its 3'-untranslated region. This leads to accumulation of *gadX* mRNA, which ultimately results in increased GAD expression (105).

One of the major repressors of GAD system is the histon-like protein, H-NS (167). H-NS directly downregulates the GadE expression and affects the GAD system (55). Another repressor of the GAD system is TorR, the response regulator of the TorS/TorR two component system, which negatively regulates *gadE*, *gadABC* and many AFI genes (18). It has been recently shown that Ler, the key positive regulator of LEE also negatively regulates GAD expression (1).

Post transcriptional regulation of GAD system is not clear. However, the fact that exponential phase cells grown under moderate acidity (pH 5.5) express high level of GAD genes but remain acid sensitive indicates the presence of regulation at translational level.

Acid fitness island (AFI)

Fitness islands are horizontally transferred genetic elements that provide advantageous traits that are not directly related to the virulence of the organism. One such element in *E. coli* and *Shigella* is the AFI, the acquisition of which is considered as a crucial step in the early evolution of these organisms. The AFI is located at 78.8 min in *E. coli* K-12 chromosome and encodes 12 genes that are involved in acid and stress resistance including *gadA* (Fig. 1.4) (8). Also encoded

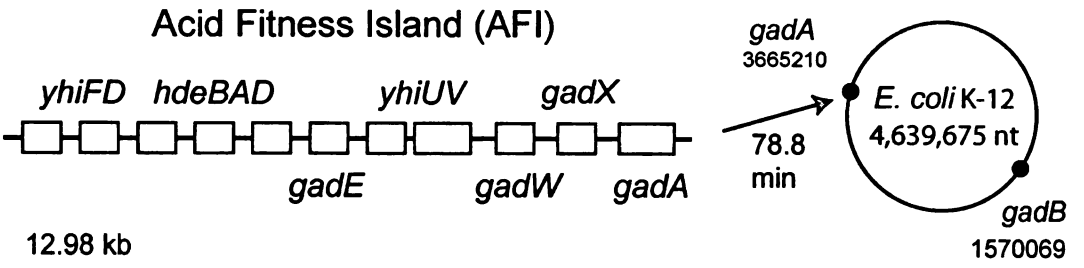


Figure 1.4. The AFI region of the *E. coli* K-12 genome. Previous studies have shown that the 12-kb AFI is located at 78.8 min on the K-12 chromosome and includes the *gadA* gene and other metabolic stress-related genes, whereas the duplicate *gadB* locus occurs across the K-12 chromosome at 33.0 min (8).

on AFI are three important regulators of the GAD system, GadE, GadX, and GadW. Many of the other AFI proteins confer different types of AR in *E. coli*. For example, the periplasmic acid stress chaperones, HdeAB, a putative LuxR family regulator, YhiF, and the lipoprotein, Slp, are required for protection against self-metabolic products at low pH and in spent medium. Similarly, a putative MgtC-family transporter, YhiD, and a predicted inner membrane protein, HdeD, are necessary for high cell density-dependent AR (71, 91). GadE regulates the expression of all the AFI genes in *E. coli* K-12 and hence, is involved in resistance to metabolic stress and high cell-density AR (91). Functions of the remaining AFI genes, *yhiU* and *yhiV* remain unclear.

GadE regulator

GadE is a LuxR family regulator with a potential helix-turn-helix DNA-binding domain in the secondary structure. GadE binds to a conserved 20 bp GAD box sequence upstream of *gadA* and *gadBC* and activates the transcription of these genes (22). A 798 bp intergenic region between *gadE* and the upstream gene *hdeD* is essential for the expression of *gadE*. The highly conserved sequence between -750 and -1 is considered as the *gadE* sensory integration region as this region coordinates the environmental and genetic regulators of *gadE* expression. At least nine activators and repressors of *gadE* converge at this region including EvgA, YdeO, GadE, TorR, Hns, PhoP, TrmE, GadX and GadW (131).

GadE is transcribed as three transcripts T1, T2 and T3. T1 starts at -124 bp, T2 at -324/-317 and T3 at -566 relative to the *gadE* translational start site with P1, P2 and P3 as corresponding promoters. P1 is the strongest promoter and is typically induced at stationary phase in complex or minimal medium. P3 is highly responsive to induction in minimal medium while P2 is expressed only in combination with P3 or alone when there is an over-expression of GadX or YdeO regulators. P2 is also activated by a TrmE-dependent manner in LB-glucose. Initial induction of GadE occurs through activation of P3P2 promoters, and once induced, GadE represses P3P2 and autoactivates P1 leading to a sustained transcription of *gadE* from P1 (131). Upon removal of inducing signals, GadE is degraded by the Lon protease, a major protease that is active in cellular protein quality control and degradation of various naturally unstable regulators (50).

GAD system and GadE in *E. coli* O157:H7

A previous study in our laboratory comparing survival rates of *E. coli* O157:H7 in complex acidic conditions such as a model stomach system (pH 2.0) with that of *E. coli* O26:H11 and *E. coli* O111:H8, revealed that *E. coli* O157:H7 has a superior ability to survive in the simulated gastric environment than the other strains tested. The quantitative PCR data from this study also showed that *E. coli* O157:H7 expresses higher transcript levels of *gadA* and *gadB* genes. This led to the suggestion that the regulation of GAD system may be different in *E. coli* O157:H7 compared to the laboratory strain *E. coli* K-12, in which most of the studies about this system have been done (10). Together, these results support

the hypothesis that the predominance of *E. coli* O157:H7 in clinical cases compared to other EHEC strains is due to its enhanced ability to resist stresses encountered in the environment both outside and inside the host. Also, this superior AR is considered a virulence factor because it contributes to the low infective dose of *E. coli* O157:H7. Furthermore, our laboratory recently showed that *gadA* and *gadB* sequences remain divergent in *E. coli* O157:H7, whereas in other *E. coli* strains, they have undergone multiple gene conversion events leading to genetic homogenization (8). All of these findings strongly suggest that the regulation and functioning of the GAD system are distinct in *E. coli* O157:H7.

Much is known about the upstream regulatory circuits and downstream effects of GadE in non-pathogenic *E. coli* (39, 84). However, GadE and its role in AR and virulence are not well characterized in any of the pathogenic *E. coli* strains. Genome comparisons between two sequenced O157:H7 strains and K12 MG1655 revealed that O157:H7 has approximately 25% additional O157-specific loci compared to *E. coli* K12 (46). During its evolution, O157:H7 has acquired many mobile elements such as lambdoid phages carrying virulence and fitness islands (121, 170). As mentioned before, important virulence factors of *E. coli* O157:H7 including shiga toxins and LEE are encoded by these horizontally transferred phage elements (77). For *E. coli*, Integrating these acquired elements into the chromosomal regulatory network is critical to becoming a successful pathogen (1). Although, the GadE sequence remains unchanged and maps to a homologous location in the O157 chromosome, it is possible that a chromosomal regulator, such as GadE, has acquired additional functions in O157:H7.

However, effects of the GadE regulator on a genome-wide scale in *E. coli* O157:H7 are still unknown.

Interaction of GAD system regulators with LEE

The relationship between AR and virulence of pathogenic *E. coli*, especially the interaction between GAD system and LEE remains largely unknown. A few studies in the past have reported that some of the GAD system regulators negatively effect the LEE expression (24, 99, 138, 157) (Fig. 1.5). In *E. coli* O157:H7, *gadE* inactivation was shown to increase the expression of LEE encoded genes *espB*, *espD* and *tir* but not *ler* (156). Hence, GadE mediated down-regulation of LEE is considered Ler-independent and the pathway through which GadE affects LEE is not known (95). Two upstream regulators of *gadE* also negatively affect expression of LEE in EPEC. One of them is GadX, which negatively regulates the expression of LEE through a plasmid encoded regulator, PerA (138). Similarly, EvgA represses the LEE expression in EPEC by activating YdeO and YdeP (99). But, the regulation of LEE by GadX and EvgA in EHEC has not been studied. It is possible that GadX down-regulates the expression of LEE through Pch, the PerC-homolog in EHEC. Additionally, because YdeO is a positive regulator of *gadE,* the down-regulation by EvgA might be mediated through GadE. However, a detailed study is warranted in this aspect in order to fully understand the mechanism of LEE repression by GAD system regulators in EHEC.

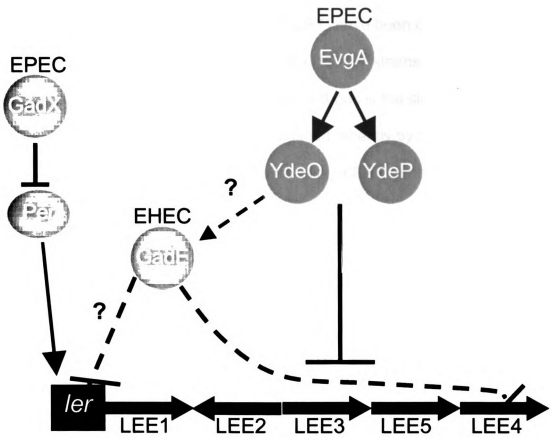
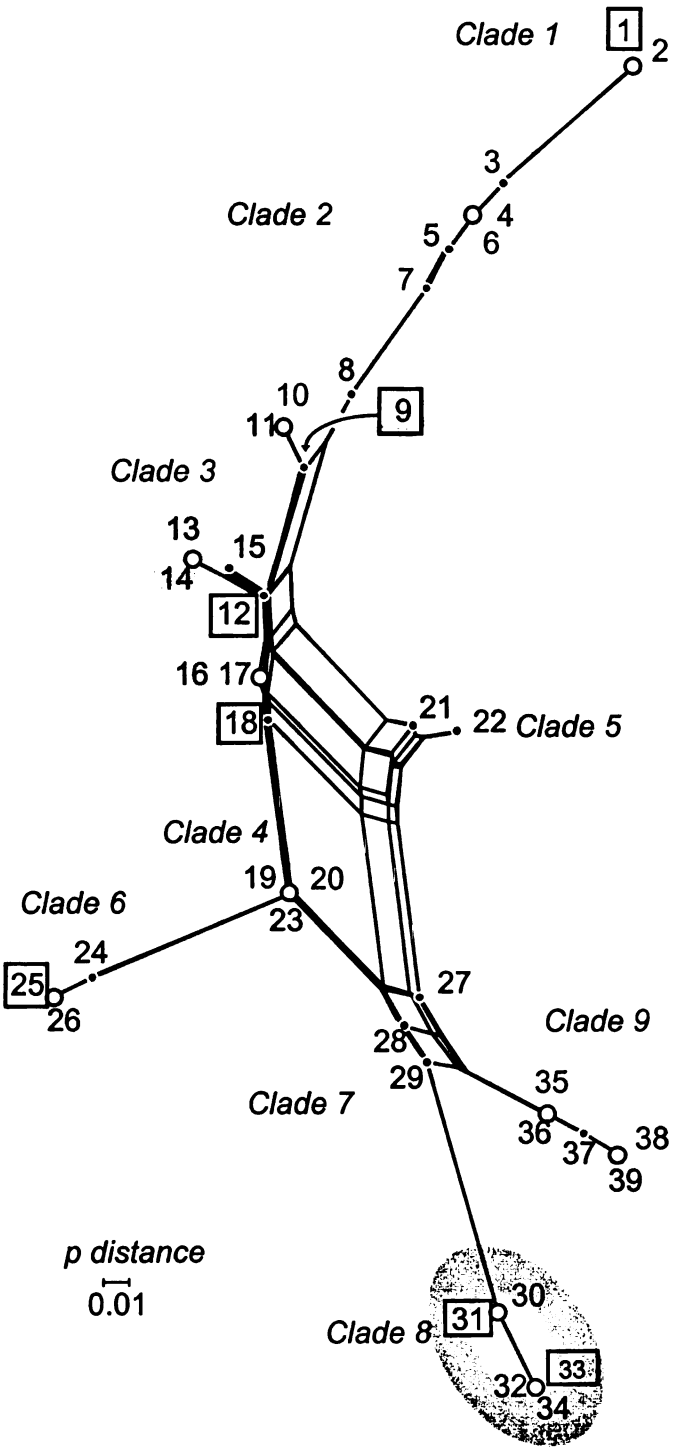


Figure 1.5. Interactions between GAD regulators and LEE. Solid lines indicate experimentally confirmed mechanisms and dotted lines indicate suggested mechanisms that are not proved experimentally.

GENOTYPES OF *E. COLI* O157:H7

Several genotyping methods such as pulse field gel electrophoresis (PFGE) and multi locus sequence typing (MLST) have been employed to investigate the genetic relatedness of *E. coli* O157:H7 strains from various sources. One of the most sensitive of these methods is the single nucleotide polymorphism (SNP) typing that was established recently by our laboratory (88). This study typed >500 clinical O157 strains using a SNP genotyping scheme that targets 96 SNP loci. Phylogenetic analysis classified the strains into 39 SNP genotypes, which formed nine distinct clades (Fig. 1.6) (88). Interestingly, the clades showed differences in the frequency and distribution of Stx genes and in the spectrum of clinical diseases reported. Importantly, clade 8 was found to be an emergent subpopulation with more chances of causing severe disease with HUS whereas groups such as clade 7 was less frequent in clinical cases and caused less severe disease. Both spinach and lettuce strains that caused fresh produce-associated outbreaks with severe disease and death in 2006 belonged to clade 8. The O157 genome strain Sakai, that was implicated in the 1996 radish sprouts outbreak in Japan belonged to clade 1 whereas the strain that caused a Hamburger outbreak in northwest regions of the United States grouped to clade 2 (88). Even though SNP typing has been conducted extensively on clinical strains of *E. coli* O157:H7, strains of bovine origin are yet to be classified by this method.

Figure 1.6. The phylogenetic network applied to 48 parsimoniously informative (PI) sites using the Neighbor-net algorithm for 528 *E. coli* O157 strains. The ellipses mark clades supported in the minimum evolution phylogeny. The numbers at the nodes denote the SNP genotypes (SGs) 1–39, and the white circle nodes contain two SGs that match at the 48 PI sites. The seven SGs found among multiple continents are marked with squares (88).


Clade 1
1
2
3
Clade 2
4
5
6
7
8
9
10
11
Clade 3
13
14
15
12
16
17
18
21
22
Clade 5
Clade 4
19
20
23
Clade 6
24
25
26
27
Clade 9
28
29
35
Clade 7
36
37
38
39
p distance
0.01
30
31
Clade 8
33
32
34

A recent study by Besser et al.(2007) classified *E. coli* O157:H7 strains of bovine (n=80) and clinical (n=282) origin into various genotypes based on the Stx-encoding bacteriophage insertion sites (14). A greater diversity of Stx-encoding bacteriophage insertion sites was observed in strains from bovine sources compared to the strains from clinical sources and these two groups were classified into different genotypes based on this method. Subsequently, strains that were predominant in clinical cases were typed as clinical genotypes whereas strains that were isolated mostly from bovine sources were considered bovine-biased genotypes (14). The basis for the variation in prevalence of clinical and bovine-biased genotypes in human clinical cases is not clear.

RATIONALE FOR THIS STUDY

E. coli O157:H7 has to survive a number of environmental stresses during transmission from the reservoir host, typically cattle, to the human gastrointestinal tract. Surviving acid stress is a critical component of transmission, as the typical human stomach pH ranges from 1.5-3.0 (112). However, *E. coli* O157:H7 has a greater average level of survival in complex acidic conditions such as the gastric environment than other groups of EHEC (10). This superior AR contributes to the low infective dose of *E. coli* O157:H7. The GAD system is the most effective AR system in *E. coli* and GadE is the central activator of the GAD system (39). Even though the GAD system and GadE are well characterized in non pathogenic laboratory strains of *E. coli*, they remain poorly defined in *E. coli* O157:H7. It is possible that a chromosomal regulator, such as GadE, has acquired additional functions in O157:H7 to integrate the mobile virulence genes acquired during its evolution into the chromosomal regulatory network. In order to successfully colonize the host, pathogenic *E. coli* must conserve energy by orchestrating their gene expression profiles in such a way that only necessary genes are expressed at each step of the infection process. It is possible that during passage through the human stomach, *E. coli* O157:H7 down-regulates the virulence mechanisms needed for later life in colon to conserve energy and activates the AR mechanisms to survive the extreme acidic pH, thus leading to a negative interaction between AR and virulence. However, the effect of acidic pH on virulence factors such as the locus of enterocyte effacement (LEE) in *E. coli* O157:H7 has not been investigated.

Based on the insertion sites of Shiga toxin-encoding bacteriophages, *E. coli* O157:H7 strains isolated from cattle and human sources are classified into clinical and bovine-biased genotypes (14). Clinical genotypes are isolated from both cattle and humans and have been shown to be associated with human disease whereas bovine-biased genotypes are isolated mostly from bovine sources (14). The varying infectivity of clinical and bovine-biased genotypes could be due to the differences in expression of virulence and stress fitness-associated genes. This proposal aims to investigate the relationship between AR and virulence of *E. coli* O157:H7 and to understand whether differences in expression of virulence and stress fitness-associated genes can explain the variation in infectivity of bovine-biased and clinical genotypes, with the following specific aims:

Specific Aim 1: To elucidate the role of GadE in acid resistance and virulence of *E. coli* O157:H7.

Hypothesis: GadE regulates the expression of the GAD system and LEE pathogenicity island in *E. coli* O157:H7, thus contributing to its acid resistance and virulence.

Specific Aim 2: To determine the effect of acidic pH on the expression of the LEE pathogenicity island and to understand the role of GAD system regulators in this effect.

Hypothesis: Acidic pH upregulates the expression of GAD system regulators, which in turn repress the expression of LEE genes.

Specific Aim 3: To determine whether the differences in infectivity of clinical and bovine-biased genotypes of *E. coli* O157:H7 is due to the differential expression of virulence and stress fitness associated genes.

Hypothesis: Virulence and stress fitness associated genes are differentially expressed between clinical and bovine-biased genotypes of *E. coli* O157:H7.

Chapter 2 describes the findings of specific aim 1 and parts of results from specific aim 2. Chapter 3 is a follow up to chapter 2 that describes remaining results from specific aim 2 and chapter 4 describes results from specific aim 3. Summary of this study and future directions are described in chapter 5.

CHAPTER 2

Characterization of the *Escherichia coli* O157:H7 Sakai GadE regulon

Kailasan Vanaja, S., T. M. Bergholz, and T. S. Whittam. 2009. Characterization of the Escherichia coli O157:H7 Sakai GadE regulon. Journal of Bacteriology 191:1868-77

SUMMARY

Integrating laterally acquired virulence genes into the backbone regulatory network is important for the pathogenesis of *Escherichia coli* O157:H7, which has captured many virulence genes through horizontal transfer during evolution. GadE is an essential transcriptional activator of the glutamate decarboxylase (GAD) system, the most efficient acid resistance (AR) mechanism in *E. coli*. The full contribution of GadE to the AR and virulence of *E. coli* O157:H7 remains largely unknown. We inactivated *gadE* in *E. coli* O157:H7 Sakai and compared global transcription profiles of mutant with that of wild type in exponential and stationary phases of growth. Inactivation of *gadE* significantly altered the expression of 60 genes independent of growth phase and 122 genes in a growth phase-dependent manner. Inactivation of *gadE* markedly down-regulated the expression of *gadA*, *gadB*, *gadC* and many acid fitness island genes. Nineteen genes encoded on the locus of enterocyte effacement (LEE), including *ler,* showed a significant increase in expression upon *gadE* inactivation. Inactivation of *ler* in Δ*gadE* reversed the effect of *gadE* deletion on LEE expression, indicating that Ler is necessary for LEE repression by GadE. GadE is also involved in down-regulation of LEE expression at moderately acidic pH. Characterization of AR of Δ*gadE* revealed that GadE is indispensable for a functional GAD system and for survival of *E. coli* O157:H7 in a simulated gastric environment. Altogether, these data indicate that GadE is critical for AR of *E. coli* O157:H7 and that it plays an important role in virulence by down-regulating expression of LEE.

INTRODUCTION

Escherichia coli O157:H7 is the prevalent variant of enterohemorrhagic *E. coli* (EHEC) associated with hemorrhagic enteritis and hemolytic uremic syndrome (HUS) in humans in the United States (3, 67). *E. coli* O157:H7 has to survive a number of environmental stresses during transmission from cows to humans. Surviving acid stress is critical during transmission, as the typical human stomach pH ranges from 1.5-3.0 (112). *E. coli* strains are more AR than other enteric pathogens and this AR is considered a virulence factor in *E. coli* O157:H7 as it contributes to the low infective dose (23, 112). *E. coli* have four distinct AR mechanisms, the oxidative (OXI) system, glutamate decarboxylase (GAD) system, arginine decarboxylase (ARG) system and lysine decarboxylase (LYS) system (39, 85) that are phenotypically distinct and provide protection against low pH dependent on the type of acidic environment encountered (119). In addition to the defined mechanisms, other factors of the general stress response, including the stress response sigma factor RpoS and the DNA binding protein Dps, also contribute to the AR of *E. coli* (28, 169).

The GAD system is the most effective system in protecting *E. coli* cells against low pH compared to other known AR mechanisms (23, 24, 75, 161). The GAD system has three components, two GAD isozymes, GadA and GadB, and the GABA-glutamate antiporter, GadC (39, 84). The *gadA* is a member of the acid fitness island (AFI), which is located at 78 min, whereas *gadB* and *gadC* form a separate operon located at 33 min in the *E. coli* K12 chromosome (55, 91,

140). Environmental signals that induce the GAD system include entry into stationary phase and acidic pH (84).

Regulation of the GAD system is complex, involving multiple regulatory circuits that influence the expression of GAD components through the central activator, GadE (39). GadE, a LuxR family regulator, is transcribed as two transcripts of sizes 0.68 kb and 1.06 kb and its secondary structure contains a potential helix-turn-helix DNA-binding domain (54, 84). However, a recent study demonstrated that GadE is possibly transcribed as three transcripts of sizes 0.9 kb, 1.1 kb and 1.38 kb (J. W. Foster and A. Sayed, presented at the 108th General Meeting of American Society of Microbiology, Boston, MA, 1 to 5 June, 2008). GadE binds to a conserved 20 bp GAD box sequence upstream of *gadA* and *gadBC* in *E. coli* K12 and activates the transcription of these genes (22, 54, 84). Although the GAD system and GadE are well-characterized in *E. coli* K12, they remain poorly defined in *E. coli* O157:H7.

A study from our lab demonstrated that *E. coli* O157:H7 strains have a greater average level of survival in complex acidic conditions, such as a simulated gastric environment, compared to other serogroups of EHEC (11). O157:H7 also expresses higher transcript levels of *gadA* and *gadB* genes than other EHEC strains in minimal medium containing glucose (11). Also, we recently showed that *gadA* and *gadB* sequences remain divergent in *E. coli* O157:H7 compared to other *E. coli* strains (8). Taken together, these findings suggest that the regulation and function of the GAD system may be distinct in *E. coli* O157:H7.

Genome sequence comparisons revealed that the O157:H7 Sakai strain has approximately 1650 O157-specific loci compared to *E. coli* K12 (47). Through evolution, the O157 population has acquired many mobile elements such as lambdoid phages carrying virulence and fitness islands (121, 171). Some of the important virulence factors of *E. coli* O157:H7, including Shiga toxins and the LEE, are encoded by horizontally transferred phage elements (78). The LEE, which is encoded on an acquired pathogenicity island, encodes a type three secretion system that mediates intimate adherence of bacteria to the intestinal mucosa through formation of attaching and effacing (AE) lesions (93). Integrating these acquired elements into the chromosomal regulatory network is critical for a pathogen to be successful (1). An example of this is one of the GAD system regulators, GadX, which has been shown to influence the expression of the LEE (138). Hence, it is possible that a chromosomal regulator such as GadE has acquired additional functions, though the effects of the GadE regulator on a genome-wide scale are still unknown for *E. coli* O157:H7. Comparison of GadE amino acid sequences among pathogenic and non pathogenic *E. coli* strains revealed no significant divergence. Recently, a study by Tatsuno et al., found that inactivation of *gadE* increases the expression of many LEE genes in O157:H7 (156). However, *gadE* inactivation did not affect the expression of the LEE encoded regulator, *ler*, and hence the pathway through which this LEE down-regulation occurs was not identified. Recently, Ler was found to negatively regulate expression of *gadE* and it was suggested that there is a reciprocal negative interaction between Ler and the GadE regulators (1).

The objectives of this study were to identify the genes regulated by GadE in *E. coli* O157:H7 and to gain insight into the mechanism underlying the negative regulation of the LEE by GadE. By comparing whole genome transcription profiles of *E. coli* O157:H7 Sakai and isogenic Δ*gadE*, we found that *gadE* positively influences expression of the GAD system genes and other AFI genes, whereas it negatively impacts the expression of the LEE genes, including *ler*. Expression of *gadE* was markedly increased in stationary phase, thereby affecting the expression of numerous genes in a growth phase-dependent manner. In addition, we also demonstrate that a functional Ler is necessary for the down-regulation of LEE by GadE. Characterization of the AR phenotype of Δ*gadE* revealed that GadE is indispensable for a functional GAD system and for survival of *E. coli* O157:H7 in a simulated gastric environment.

MATERIALS AND METHODS

Bacterial strains and growth conditions. Bacterial strains and plasmids used in this study are summarized in Table 2.1. All strains were stored at -70°C in LB broth containing 10% glycerol, inoculated into 10 ml LB broth, and grown to an OD_{600} of ~0.1 to recover cells. To minimize the confounding effect of acidic pH that would develop in stationary phase of growth in un-buffered medium, the strains were grown in MOPS minimal medium buffered to pH 7.4. Cells recovered in LB were then grown twice to stationary phase in MOPS-buffered minimal medium before a final transfer at 1:30 dilution into 100 ml MOPS medium for RNA isolation and model stomach assay (12).

Genetic manipulations. *E. coli* O157:H7 Sakai Δ*gadE* and Δ*ler* strains were constructed by the modified one-step gene inactivation method for EHEC developed by Murphy et al (32, 98). Briefly, recombinant PCR products containing a kanamycin (Km) resistance marker flanked by 45-50 bp sequences homologous to the upstream and downstream regions of target genes were generated using the primers listed in Table 2.2 from plasmid pKD4 (32). PCR products were electroporated into red recombinase-producing *E. coli* O157:H7 (TW15901) as described (98) and the transformants were identified on LB agar plates with 25 μg/ml Km at 37°C. The Km resistance marker was removed from the Δ*gadE* by introducing plasmid pCP20 that encodes FLP recombinase (32). Subsequently, the double mutant, Δ*gadE*Δ*ler* was constructed by one step inactivation of *ler* in Δ*gadE* using the method described above.

TABLE 2.1

Bacterial strains and plasmids used in this study

Strain or plasmid	Genotype	Reference or Source
Strains		
TW08264	*E. coli* O157:H7 RIMD0509952 (Sakai) wild type	(96)
TW15901	TW08264 harboring pKM208 plasmid	This study
TW15902	Δ*gadE*	This study
TW15903	Δ*gadE*Δ*ler*::Km	This study
TW15904	Δ*ler*::Km	This study
TW15905	Δ*gadE*/pCR2.1*gadE*	This study
Plasmids		
pKM208	Red recombinase expression plasmid, Ap^r	(32)
pKD4	Template plasmid for lambda Red recombination system, Km^r	(32)
pCP20	Flp recombinase expression plasmid, Cm^r	(32)
pCR2.1	Cloning vector	Invitrogen
pCR2.1*gadE*	*gadE* cloned into pCR2.1	This study

TABLE 2.2

Oligonucleotide primers used for one step inactivation

Primer[a]	Sequence (5' to 3')[b]
gadE-H-F	ATCAATTCCCTGTCAGAGATCAAAAAAGTAGGCAATAAACCCTTCAAGG Tgtgtaggctggagctgcttc
gadE-H-R	CTCGTCATGCCAGCCATCAATTTCAGTTGCTTATGTCCTGACTAAAAATA catatgaatatcctccttag
ler-H-F*	TTTCATCTTCCAGCTCAGTTATCGTTATCATTTAATTATTTCATGgtgtaggct ggagctgcttc
ler-H-R*	GTTGGTCCTTCCTGATAAGGTCGCTAATAGCTTAAAATATTAAAGcatatga atatcctccttag

[a]The homology regions for the *ler* primers are from Iyoda et al., (60)

[b] Priming sites for pKD4 are in lowercase letters

For complementation of the Δ*gadE* strain, DNA fragments of 2,630 bp containing the Sakai *gadE* coding region and additional flanking regions of *gadE* were amplified from *E. coli* O157:H7 Sakai chromosomal DNA using *TaKaRa LA Taq*™ polymerase (Takara Bio USA, Madison, WI). The resulting PCR products were cloned into pCR®2.1-TOPO® vector (Invitrogen, Carlsbad, CA) to make pCR2.1*gadE* plasmid, which was transformed into the Δ*gadE* strain creating the Δ*gadE*/pCR2.1*gadE* strain.

RNA isolation and cDNA labeling. For RNA isolation, the wild type Sakai and Δ*gadE* strains were grown to early exponential phase (2.25 h, OD_{600} ~0.25) and stationary phase (5.5 h, OD_{600} ~1.5) in MOPS medium as described above. RNA was isolated from five independent cultures using hot acid-phenol:chloroform extraction. At each growth phase, 5 ml of culture was mixed with equal volume of hot acid-phenol:chloroform (pH 4.5 with Iso Amyl Alcohol (IAA), 125:25:1) (Ambion, Austin, TX) and incubated at 65°C with periodic shaking for 10 min. The samples were centrifuged at 3220 × g for 20 min and the supernatant was subjected to further extractions with phenol:chloroform and chloroform:IAA (12). RNA was precipitated overnight at -70°C in 2.5 volume 100% ethanol and 1/10 volume 3 M sodium acetate pH 5.2. RNA purification and DNase treatment of RNA samples were done with the RNeasy kit (Qiagen, Valencia, CA) and RNA quality was assessed on a formaldehyde-agarose gel.

Six µg of RNA was used for reverse transcription reactions containing 3 µg random primers (Invitrogen), 1x first strand buffer (Invitrogen), 10 mM DTT, 400 U Superscript II (Invitrogen), 0.5 mM of dATP, dCTP and dGTP, 0.3 mM

dTTP and 0.2 mM amino-allyl dUTP (12). After incubation at 42°C overnight, the cDNAs were purified with Qiagen PCR clean up columns with phosphate wash buffer (5 mM K_2HPO_4, pH 8.0, 80% ethyl alcohol) and phosphate elution buffer (4 mM K_2HPO_4, pH 8.5) and were coupled with either Cy3 or Cy5 dyes (Amersham Biosciences, Piscataway, NJ) as described previously (12).

cDNA hybridizations. Hybridizations were performed according to a loop design, which included between strain (wild type vs. mutant at same growth phase) and within strain (same strain at two growth phases) comparisons. Five biological replicates were included for each comparison resulting in 20 arrays. As described (12), the cDNAs were hybridized onto microarray slides printed with 6088 ORFs including 110 ORFs from pO157 plasmid, and representing *E. coli* strains K12, EDL933 and Sakai. Arrays were scanned with an Axon 4000b scanner (Molecular Devices, Sunnyvale, CA) followed by image analysis using GenePix 6.0 (Molecular Devices) (12).

Data analysis. The microarray data were analyzed using R (v. 2.2.1) and the MAANOVA (v. 0.98.8) package. Raw intensity values from replicate probes were averaged and $\log_2$ transformed after normalization with the pin-tip LOWESS method. The normalized intensity values were fitted to a mixed model ANOVA considering array and biological replicates as random factors and dye, strain and growth phase as fixed factors (30). The linear model tested was Y (intensity) = array + dye + strain (wild type or mutant) + growth phase (exponential or stationary) + strain*growth phase + sample (biological replicate) + error. Each main effect had 2 levels: mutant and wild type for strain and exponential and

stationary phases for growth phase. The design included between and within strain comparisons using 5 biological replicates. Significant differences in expression due to strain, growth phase and strain*growth phase were determined using the Fs test in MAANOVA, which uses a shrinkage estimator for gene-specific variance components that makes no assumption about the variances across genes (31) with 500 random permutations to estimate the p-values. ANOVA with a mixed linear models have been used to analyze microarray experiments with repeated measures where transcript levels of the strains at two different growth phases are measured (6, 63, 80). The q-value package in R was used for determining the false discovery rate (FDR) (146).

Overrepresentation of gene sets with a common biological function in the wild type or mutant strain were determined using the GSEA Preranked analysis in Gene Set Enrichment Analysis (GSEA v2.0) program (Broad Institute, Massachusetts Institute of Technology) (148). The gene sets were designated based on the TIGR annotation for the Sakai genome (http://cmr.jcvi.org/tigr-scripts/CMR/GenomePage.cgi?org=ntec03). Additionally, two gene sets, the LEE and the AFI and GAD, also were included in the analysis (33, 41, 91, 158). The pattern search tool in *coli*BASE (http://xbase.bham.ac.uk/colibase/pattern.pl?id=1073) was used to identify GAD box (22, 85) sequences (5'-TTAGGATTTTGTTATTTAAA-3') in the putative promoter regions of genes differentially regulated between wild type and Δ*gadE*. A cut off of 70% similarity to the query sequence was set to apply higher stringency because experimental confirmation of GadE binding was not

conducted. A sequence logo for the consensus sequence was created at http://weblogo.berkeley.edu/logo.cgi.

Quantitative real-time PCR. RNA isolations were conducted at both exponential and stationary phases of growth. For assays with Δ*gadE*/pCR2.1*gadE*, Δ*gadE*Δ*ler* and Δ*ler*, RNA was isolated only at the exponential phase. Taqman assays (11) were used for quantifying the expression of *gadA*, *gadB* and *ler* with *mdh* as reference for normalization. For the remaining genes, SYBR green chemistry was used for measuring expression levels. Primers were designed using the Primer3 server (125) based on the published reference genome sequence of *E. coli* O157:H7 strain Sakai (Table 2.S3 in appendix). cDNA synthesis was conducted using iScript Select cDNA synthesis kit (BioRad, Hercules, CA) with 1 µg of total RNA according to the manufacturer's instructions. After reverse transcription, 5-fold serial cDNA dilutions were used for Q-PCR assays containing 12.5 µl 2x iQ SYBR green supermix (BioRad), 0.63 µl of each primer (10 µM stock), 9.24 µl of H2O and 2 µl cDNA with cycle conditions of 95°C for 2 min followed by 40 cycles of 10 sec at 95°C and then 20 sec at the specific annealing temperature (12). The expression levels of the 16S rRNA gene were used for normalization of data and the relative expression levels were quantified using Pfaffl's method (113). The results presented are averages from at least three biological replicates ± standard error of mean (SEM).

Expression studies in EG minimal medium. Wild type and Δ*gadE* cells were grown in EG minimal medium at pH 7.0 and pH 5.0 (83) to late exponential

phase (OD_{600} ~0.5). RNA extractions were conducted using a modified hot-phenol extraction protocol (15) that utilized 5% acidic phenol in ethanol. cDNA synthesis and Q-PCR methods are described above.

AR mechanism assays. AR mechanism assays for the GAD, ARG and OXI systems were conducted as described previously (75, 83). Briefly, for the GAD system, strains grown in LB broth with 0.4% glucose (LBG) were challenged at pH 2.0 in a test environment (EG + glutamate) and in a control environment (EG), whereas for the ARG system, after growth in BHI broth with 0.6% glucose, strains were tested in test (EG + arginine) and control (EG) environments at pH 2.5. For testing the OXI system, strains were grown in LBMES (pH 5.0) and EG (pH 7.0) and challenged at pH 2.5 in EG. Samples were withdrawn at specific time points (30 min or 1 h intervals) and plated on LB agar plates using an Autoplate 4000 Spiral Plater (Spiral Biotech, Bethesda, MD). Colonies were counted after overnight incubation at 37°C using the Q-Count (Spiral Biotech). Assays were conducted for at least two biological replicates, each with two technical replicates. CFU/ml from technical replicates were averaged and converted to log10 CFU/ml. The results reported are averages ± SEM for at least three experiments.

Model stomach assay. The model stomach system (MSS) (64) was prepared as described previously (11). Gerber Turkey Rice Dinner© baby food (30 g) was mixed with 120 ml of synthetic gastric fluid (pH 1.75) yielding a final acidity of pH 2.5. Contents of the MSS were stomached for 30 sec, sampled, diluted, and plated onto LB agar plates every 30 min for 1.5 h to enumerate

viable cells. CFU/ml from duplicate plates were averaged and converted to log10 CFU/ml. The results reported are averages ± SEM for three experiments.

Microarray data accession. Microarray data are available at NCBI GEO (http://www.ncbi.nlm.nih.gov/geo), accession number GSE13132.

RESULTS

Identification of genes regulated by GadE. Inactivation of *gadE* did not cause a significant difference in the growth rate of *E. coli* O157:H7; the generation time of wild type was 43.7 min and that of Δ*gadE* was 44.1 min. A two-factor ANOVA with two main effects (strain and growth phase) and the interaction effect (strain*growth phase) was used to determine the impact of *gadE* inactivation on the transcriptome of *E. coli* O157:H7 at both exponential and stationary phase. Genes with FDR < 0.1 for strain effect and FDR < 0.05 for interaction effect were considered as regulated by GadE. Significant strain effects were identified for 60 genes, indicating differential expression between the wild type and Δ*gadE* (FDR < 0.1) (Table 2.S1 in appendix). Of these, 58 genes had higher transcript levels in Δ*gadE* demonstrating that GadE has a negative influence on their transcription, and 2 genes had lower transcript levels in Δ*gadE,* indicating that GadE has a positive influence on their transcription. Among the 2 genes with lower transcript levels in Δ*gadE*, ECs3904 had significantly greater transcript levels in exponential phase compared to stationary phase, and ECs2294 had greater levels in stationary phase. Among the 58 genes with higher transcript levels in Δ*gadE*, 33 were significantly higher in exponential phase, including LEE genes *tir*, *espF* and *cesT*, and 25 were higher in stationary phase, including genes *vgrE*, *ilvG*, and *treR* (Table 2.S1 in appendix). A significant interaction effect, which indicates that inactivation of *gadE* affects expression of genes differently at each growth phase, was identified for 122 genes, including the AFI and GAD genes *gadA, gadB* and *gadC* (FDR <

0.05) (Table 2.S2 in appendix). GSEA analysis identified significant enrichment of the AFI and GAD genes (FDR < 0.05) in wild type and LEE genes (FDR < 0.05) in Δ*gadE*. In summary, the array data demonstrated that inactivation of *gadE* had an effect on several genes, including members of the LEE pathogenicity island, in addition to the GAD and AFI genes.

Interaction effects of *gadE* inactivation and growth phase.

Expression of *gadE* was 84.2 ± 13.4-fold higher in stationary phase compared to exponential phase in wild type as measured by Q-PCR. Since the expression of *gadE* is growth phase-dependent, it affected the expression of 122 genes in a growth phase-dependent manner leading to a significant interaction effect (FDR <0.05) (Table 2.S2 in appendix). Genes with a significant interaction effect included several genes belonging to the AFI and GAD system, as well as a number of genes involved in energy metabolism and genes encoding transcriptional regulators. In exponential phase, GadE exhibited a positive effect on transcript levels of genes such as *cyoDC, sdhCDAB,* and *sucAB* involved in energy metabolism, while in stationary phase, GadE exhibited a negative effect on transcript levels of these genes (Table 2.S2 in appendix). Transcript levels of genes encoding sigma factors and transcriptional regulators such as *rpoS, lysR,* and *pspF* were slightly elevated in Δ*gadE* in exponential phase, but were 1.3- to 2.9-fold higher in the wild type compared to Δ*gadE* in stationary phase, indicating that these genes are positively regulated by GadE at stationary phase (Table 2.S2 in appendix).

Putative GadE binding site upstream of GadE regulated genes. To determine whether differentially expressed genes identified by the microarray analysis are regulated directly by GadE, a pattern search against the *E. coli* O157:H7 Sakai genome was conducted in *coli*BASE to identify potential GAD boxes upstream of these genes. A conserved GAD box sequence described previously in *E. coli* K12 (22, 85) was used for the pattern search. Typically one GAD box is observed upstream of GadE regulated genes, *gadA* and *gadBC*, in *E. coli* K12 (22). Sequences with ≥70% similarity to the query sequence were considered as putative binding sites of GadE. Matching sequences were detected upstream of 8 genes that were significantly differentially expressed in Δ*gadE*. As expected, GAD box sequences preceding *gadA* and *gadB* showed 100% similarity to the conserved sequence in *E. coli* K12. There were two putative GadE binding regions identified upstream of *hdeD*. Matching sequences also were identified in the upstream regions of *vgrE* and 3 LEE genes (*sepZ*, *escC*, and *ler*) (Fig. 2.1).

Expression of AFI and GAD genes in Δ*gadE*. Inactivation of *gadE* resulted in a decrease in expression of many of the AFI and GAD genes and the magnitude of decrease was dependent on the growth phase. Six AFI and GAD genes, including *gadA, gadBC* and *hdeBAD* had a significant interaction effect (Table 2.3). These results were verified with Q-PCR. At exponential phase, for *gadA*, *gadB*, *gadC*, *hdeA* and *hdeB*, microarray data revealed higher expression in the mutant, whereas Q-PCR detected higher expression in wild type. This discrepancy could be due to the negligible expression of these genes at

Gene	% identity	Sequence
gadA	100	TTAGGATTTTGTTATTTAAA
gadB	100	TTAGGATTTTGTTATTTAAA
pspF	70	TTA**TCTTTTT**GA**TT**TATAAA
*hdeD1**	75	TTAGGAAA**TTTTT**ATTAAAT
*hdeD2**	70	**ATCAGATATTTTT**ATTTCAA
vgrE	70	TGA**TT**ATTTTGTT**GA**CTAAA
sepZ	70	**GGATAATTTGGTT**ATTTATA
escC	70	TTCGA**CTCTTTTT**AATTAAA
ler	70	ATATGATTTTTTTGTTGACA

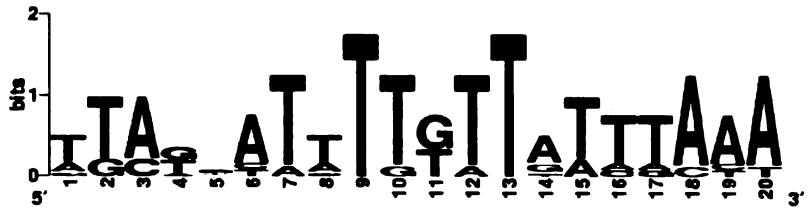


Figure 2.1. Alignment of putative GAD box sequences. Sequences upstream of 8 *gadE*- regulated genes with greater than 70% identity to the conserved GAD box sequence were identified by the pattern search tool in *coli*BASE. Letters in boldface in the table are unmatched bases. Asterisks indicate the two matching sequences found upstream of *hdeD*.

TABLE 2.3

Effect of *gadE* inactivation on expression of AFI and GAD genes in exponential and stationary phase

ECs number	Gene[a]	Function	Exponential phase expression ratio (WT/ Δ*gadE*)		Stationary phase expression ratio (WT/ Δ*gadE*)	
			Micro array	QPCR[b]	Micro array	QPCR[b]
ECs 4389	***hdeB***	Periplasmic chaperone	0.6	5.2±2.0	2.8	31.4±1.5
ECs 4390	***hdeA***	Protection from organic acid metabolites	0.3	3.2±0.3	11.1	39.3±3.0
ECs 4391	***hdeD***	Acid resistance at high cell densities	0.6		2.5	
ECs 4397	***gadA***	Glutamate decarboxylase isozyme	0.4	2.2±0.8	6.4	46.3±13.7
ECs 2097	*gadC**	Glutamate-GABA antiporter	0.5	2.3±0.5	4.1	16±0.9
ECs 2098	***gadB****	Glutamate decarboxylase isozyme	0.4	2.2±1.5	8.0	22.5±1.67

[a] Genes marked with an asterisk are GAD system genes not encoded in AFI, genes in bold face have putative GAD boxes upstream of their sequence.

[b] Fold change±SEM as determined by Q-PCR

exponential phase in minimal medium at neutral pH (84). It has been shown that microarrays are less sensitive than Q-PCR in detecting changes in expression when the transcript levels are low (21). At stationary phase, Q-PCR supported the

microarray results for these 5 genes, though a greater difference was detected in transcript levels compared to the microarray (Table 2.3). This underestimation of fold changes by microarrays has been reported in many of the previous studies (27, 174). The fold increase in expression of *gadA* and *gadB* in wild type compared to Δ*gadE* was approximately 10-20 times higher at stationary phase than at exponential phase. Similarly, the increase in expression of *hdeAB* in wild type was 6-12 times higher at stationary phase than at exponential phase (Table 2.3). In summary, both microarray and QPCR data demonstrated that the difference in expression of AFI and GAD genes between wild type and Δ*gadE* was minimal at exponential phase whereas at stationary phase there was a marked decrease in expression of *gadABC* and *hdeAB* in the mutant. Moreover, in the wild type, the expression of AFI and GAD genes increased markedly from exponential to stationary phase whereas in the Δ*gadE* their expression decreased minimally or remained unchanged as the cells entered stationary phase. This demonstrates that inactivation of *gadE* abrogates the growth phase regulation of AFI and GAD genes in *E. coli* O157:H7.

Six AFI genes did not show differential expression between the mutant and wild type, including two AraC-like regulators of the GAD system, *gadX* and *gadW*. A significant increase in expression of multi drug resistance-related efflux

pump gene, *yhiU,* was observed in the mutant (Table 2.S1 in appendix). Consistent with previous observations (12), all of the AFI and GAD genes had significantly higher expression at stationary phase than at exponential phase. GSEA analysis confirmed the enrichment of the gene set representing AFI and GAD genes in the wild type.

GadE represses expression of LEE. Inactivation of *gadE* significantly elevated the expression of 19 LEE genes independent of the growth phase, among which 8 genes showed ≥1.35-fold increase in expression in the mutant (Table 4). Fold changes in expression between mutant and wild type detected by microarray and Q-PCR were highly correlated for the LEE genes (Table 2.4). Contrary to previous findings (156), the LEE encoded regulator, *ler,* was up-regulated by 2.2-fold in the mutant as determined by Q-PCR (1.35-fold in the microarray). Other LEE genes such as *tir* and *espD* were up-regulated by 1.8- and 1.6-fold, respectively (microarray detected 1.4-fold for both) (Table 2.4). Most of the LEE genes with significant differential expression between Δ*gadE* and the wild type also had higher expression in exponential phase compared to stationary phase. Six LEE genes, including *sepZ* (*espZ*), had higher expression at stationary phase and *sepZ* showed a 1.4-fold (1.36-fold in microarray) increase in expression in the mutant. Two non-LEE encoded effectors, *nleG2-2* and *nleG2-3*, also had significant increase in expression in the mutant at stationary phase. In the array data, *eae, espA* and *espB* were not significantly up-regulated, whereas Q-PCR identified an increase in expression of 1.3-1.6-fold

TABLE 2.4

LEE genes significantly up-regulated in Δ*gadE* (FDR<0.1)

ECs number	Gene[a]	Function	Micro array (Δ*gadE* /WT)	Q-PCR (Δ*gadE*/ WT)[b]	Microarray (Stationary/ Exponential) [c]
ECs 4550	*espF*	type III secretion system, secreted effector	1.30	1.34±0.08	0.6
ECs 4553	*cesD2*	type III secretion system chaperone	1.36		0.5
ECs 4554	*espB**	type III secretion system, secreted translocator	1.20	1.38±0.09	
ECs 4555	*espD*	type III secretion system, secreted translocator	1.38	1.63±0.04	
ECs 4555	*espA**	type III secretion system, secreted translocator	1.10	1.58±0.2	
ECs 4558	*escD*	type III secretion system, structural protein	1.18		1.3
ECs 4558	*eae**	gamma intimin	1.20	1.3±0.1	
ECs 4560	*cesT*	type III secretion system, chaperone	1.38		0.3
ECs 4561	*tir*	translocated intimin receptor protein	1.39	1.84±0.06	0.5
ECs 4562	*map*	type III secretion system, secreted effector	1.26		
ECs 4563	*cesF*	type III secretion system, chaperone	1.31		1.6
ECs 4564	*espH*	type III secretion system, secreted effector	1.33		3.1
ECs 4565	*sepQ*	type III secretion system, structural protein	1.27		2.9

Table 2.4 continued...

ECs 4567	*orf15*	orf of unknown function	1.19		0.9
ECs 4571	***sepZ***	type III secretion system, secreted effector	1.36	1.44±0.06	6.4
ECs 4572	*rorf8*	orf of unknown function	1.26		3.0
ECs 4575	***escC***	type III secretion system, structural protein	1.21		0.3
ECs 4584	*orf5*	orf of unknown function	1.34		0.5
ECs 4585	*orf4*	orf of unknown function	1.41		0.6
ECs 4586	*orf3*	orf of unknown function	1.37		0.6
ECs 4587	*cesAB*	type III secretion system, chaperone	1.35		0.6
ECs 4588	***ler***	type III secretion system, regulator	1.35	2.23±0.26	0.8
ECs 4590	*espG*	type III secretion system, secreted effector	1.27		0.6
ECs 1994	*nleG2-2*	non LEE-encoded effector	1.26		2.5
ECs 2156	*nleG2-3*	non LEE-encoded effector	1.28		2.9

[a] Genes underlined are non LEE-encoded effectors and genes with asterisks are not significant in microarray, but detected as up-regulated by Q-PCR. Genes in bold face have putative GAD boxes upstream of their sequence.

[b] Exponential phase fold change±SEM as determined by Q-PCR

[c] Ratios are reported only for genes with a significant growth phase effect (FDR<0.05)

in the mutant, similar to other significant LEE genes (Table 2.4). Significant enrichment of the LEE gene set in Δ*gadE* was detected by GSEA. To further confirm the negative influence of GadE on LEE expression, exponential phase transcript levels of select LEE genes were measured in the complement strain, Δ*gadE*/pCR2.1*gadE*, which over-expresses *gadE*. Expression of *ler* decreased by 6-fold in the complement, whereas *espD* and *sepZ* decreased by 9.9- and 9.5-fold, respectively. Altogether, these data demonstrate that GadE is a repressor of LEE genes, including *ler*, in *E. coli* O157:H7.

Repression of LEE by GadE is mediated through Ler. Expression data from Δ*gadE* and *gadE*-over expressing strains demonstrated that GadE negatively regulates the expression of *ler*. Moreover, a pattern search in *coli*BASE to find GAD box sequences in the LEE island region revealed a putative GAD box upstream of *ler* (-199 to -180bp) with 6 mismatches (70% identity). To determine whether repression of LEE genes by GadE is mediated by Ler, we inactivated *ler* in both the Δ*gadE* and wild type strains and compared the expression of select LEE genes. If GadE down-regulates LEE expression independent of *ler*, then an increase in expression of LEE genes in the double mutant, Δ*gadE*Δ*ler*, similar to Δ*gadE* was expected. In this study, however, expression of *tir*, *sepZ*, *espA*, *espB* and *espD* decreased by 11.6-, 20.1-, 38.2-, 17.9- and 33.4-fold, respectively in Δ*gadE*Δ*ler* (Fig. 2.2). A similar decrease in LEE expression was observed in Δ*ler*. These data demonstrate that the positive effect of *gadE* inactivation on LEE expression is reversed by *ler* inactivation, suggesting that *ler* is essential for the repression of LEE by GadE.

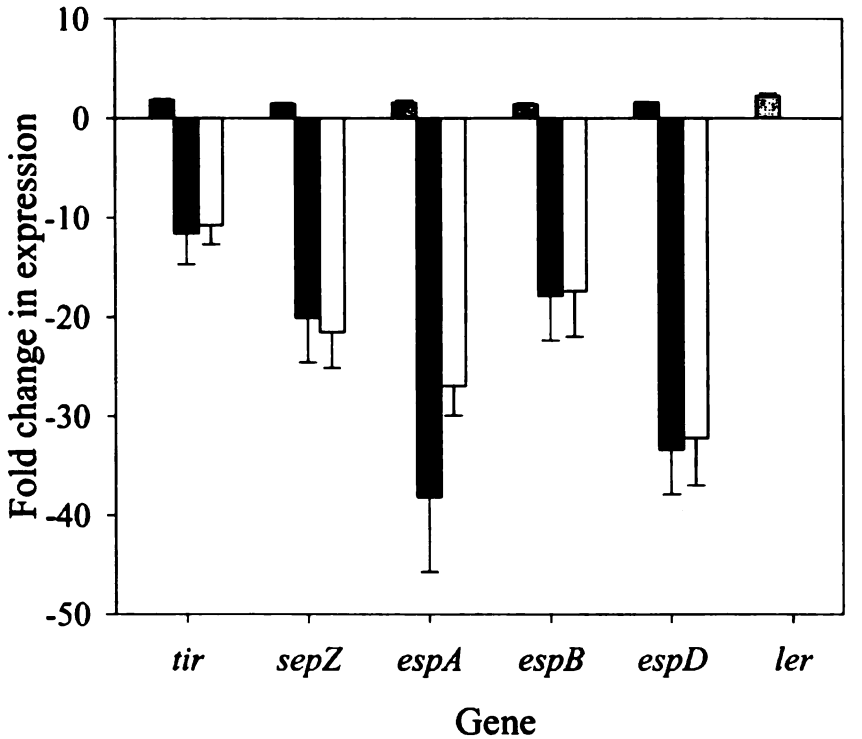


Figure 2.2. Exponential phase expression of LEE genes, *tir*, *sepZ*. *espA*, *espB*, *espD* and *ler* in Δ*gadE* (grey bars), Δ*gadE*Δ*ler*::Km (black bars) and Δ*ler*::Km (white bars) compared to the wild type. Results shown are average fold change in expression measured by QPCR with standard error of mean (SEM) from at least three biological replicates.

Effect of acidic pH on expression of LEE genes. To determine the influence of acidic pH on expression of the LEE genes, wild type and Δ*gadE* cultures were grown to exponential phase (OD_{600}~0.5) in EG minimal medium adjusted to pH 7.0 (control) and pH 5.0 (moderately acidic) for comparing *gadE* and LEE expression. Growth in EG pH 5.0 resulted in a 34-fold increase in the wild type expression of *gadE*. Expression of three LEE genes, *ler*, *espD* and *sepZ*, was down-regulated in EG pH 5.0 compared to EG pH 7.0 in the wild type (Table 2.5). To determine if the down-regulation of LEE genes in response to moderate acidity was directed exclusively by GadE, the expression of LEE genes in Δ*gadE* grown in EG pH 5.0 was assessed. There was a 3.9- and 6.5-fold decrease in expression of *espD* and *sepZ*, respectively, in Δ*gadE*, however, this decrease was lower relative to the wild type, where expression of *espD* and *sepZ* decreased by 7.5- and 9.4-fold, respectively. Interestingly, the pattern of *ler* expression at acidic pH was different from the other 2 LEE genes tested. In the wild type, there was a 6-fold decrease in *ler* expression, whereas in Δ*gadE*, *ler* expression increased 4.5-fold at pH 5.0 (Table 2.5). Together, these observations indicate that down-regulation of the major LEE regulator, *ler*, is mediated through GadE in response to moderate acid stress, whereas repression of other LEE genes under the same conditions involve additional GadE/Ler-independent factors.

Two regulators that may affect the expression of LEE genes at acidic pH, independent of GadE and Ler, are GadX and EvgA, which have a negative effect on LEE expression in EPEC (99, 138). Hence, we measured the expression of

TABLE 2.5

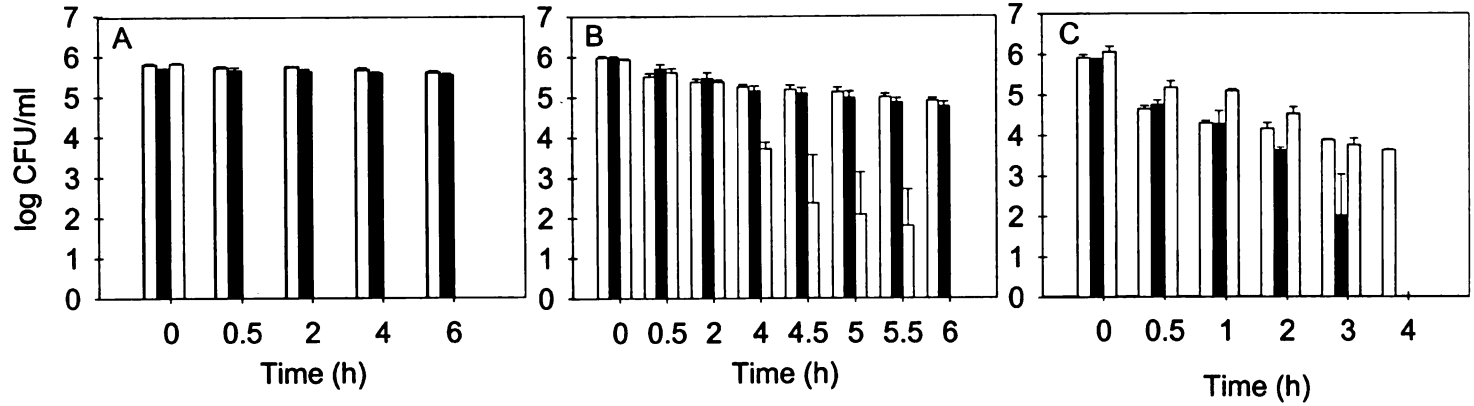
Effect of pH 5.0 on expression of GAD regulators and LEE genes

Gene	Fold change in expression (pH 5.0/pH7.0)	
	WT	Δ*gadE*
gadE	34.4±6.9	no expression
gadX	10.4±2.5	4.3±1.0
evgA	2.8±0.3	3.4±0.7
ler	0.16±0.0	4.5±0.4
espD	0.16±0.06	0.26±0.02
sepZ	0.12±0.04	0.17±0.05

these regulators in EG pH 7.0 and EG pH 5.0 in wild type and Δ*gadE* and observed a strong induction of both genes in both strains at pH 5.0. The *gadX* gene had > 2-fold higher induction in wild type, whereas *evgA* induction was similar in both wild type and Δ*gadE* (Table 2.5).

Functional GadE is necessary for optimal performance of the 3 principal AR mechanisms in *E. coli* O157:H7. To functionally confirm the microarray data, which revealed a marked decrease in expression of GAD and AFI genes in Δ*gadE,* we conducted AR mechanism assays for the GAD, ARG and OXI systems. The ARG and OXI systems were included in the study since GadE has been shown to influence their function in *E. coli* K12 (84). The Δ*gadE* strain could not survive in the test environment for the GAD system (pH 2.0 with glutamate) even for 30 min, indicating a non-functional GAD system (Fig. 2.3A). The wild type and complement showed a log reduction of 0.20 ± 0.08 and 0.17 ± 0.02 CFU/ml, respectively, after 6 h of exposure to the test environment. In the ARG system test environment (pH 2.5 with arginine), survival of Δ*gadE* was similar to the wild type and complement for up to 2 h. However, at 4 h there was reduction in viable cell numbers and the mutant showed high variation in cell numbers up to 5.5 h, and by 6 h, no viable mutants were recovered. The wild type and complement showed a log reduction of 1.07 ± 0.08 and 1.22 ± 0.16 CFU/ml, respectively, after 6 h (Fig. 2.3B). The OXI system was less effective in protecting all three strains compared to the ARG system. The mutant survived for only 3 h at pH 2.5 and the log reduction in CFU/ml after 4 h was 2.27 ± 0.1 for wild type (Fig. 2.3C). Interestingly, the complement also did not survive after 3 h,

Figure 2.3 A-C. AR mechanism assays. Survival of wild type (white bars), Δ*gadE* (grey bars) and Δ*gadE*/pCR2.1*gadE* (black bars) strains in the 3 AR systems. (A) Survival for the GAD system test at pH 2.0 (B) Survival for the ARG system test at pH 2.5 (C) Survival for the OXI system test at pH 2.5. The results presented are average CFU/ml with SEM from 3 experiments for each AR system.


A
B
C
log CFU/ml
0 1 2 3 4 5 6 7
0 0.5 2 4 6
0 0.5 2 4 4.5 5 5.5 6
0 0.5 1 2 3 4
Time (h)
Time (h)
Time (h)

indicating that *gadE* in trans does not reconstitute the phenotype for the OXI system. It is possible that flooding the cell with multiple copies of *gadE,* as in the complement, adversely effected the functioning of OXI system. These findings demonstrate that inactivation of *gadE* abolished the functioning of the GAD system and rendered ARG and OXI systems less effective in protecting the cells against low pH.

Survival of Δ*gadE* in a simulated gastric environment. Since the 3 principal AR systems were defective in protecting the Δ*gadE* from acidic stress in defined minimal test conditions, the ability of Δ*gadE* to survive in a complex acidic environment was assessed using the MSS (pH 2.5). The wild type and complemented cells showed an average log reduction of 1.05 ± 0.06 and 0.32 ± 0.04 CFU/ml, respectively, after 1.5 h in the MSS (Fig. 2.4). Viable cells could not be recovered from the MSS inoculated with Δ*gadE*, which indicates that functional *gadE* is necessary for survival in the simulated gastric environment.

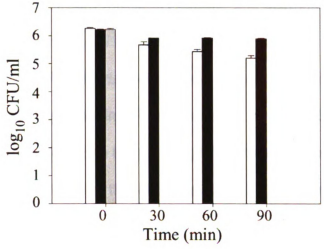


Figure 2.4. Survival of wild type (*white* bars), Δ*gadE* (grey bars) and Δ*gadE*/pCR2.1*gadE* (black bars) strains in MSS. The average log CFU/ml with SEM from 3 experiments is plotted for each time point.

DISCUSSION

Although the upstream regulatory circuits and downstream effects of GadE in non-pathogenic *E. coli* has been examined (39, 84), little is known about GadE and its role in AR and virulence among pathogenic *E. coli* strains. Here, the role of the GadE regulator in AR and virulence of *E. coli* O157:H7 was investigated by constructing an isogenic Δ*gadE* strain and comparing its expression profiles with that of the wild type strain. Our findings demonstrate that besides being a positive regulator of GAD and many AFI genes, GadE acts as a negative regulator of the LEE pathogenicity island, an important factor in the virulence of *E. coli* O157:H7. GadE, along with additional regulators, is involved in the down-regulation of LEE expression at moderately acidic pH. In addition, the characterization of AR phenotypes of Δ*gadE* revealed that GadE is indispensable for a functional GAD system and plays a vital role in the survival of *E. coli* O157:H7 in a simulated gastric environment.

In this study, the microarray data demonstrated that inactivation of *gadE* in *E. coli* O157:H7 altered expression of 60 genes independent of growth phase and 122 genes in a growth phase-dependent manner. The genes with altered expression included both AR and virulence genes, indicating that the regulatory function of GadE is not restricted to AR, but has a more global effect on the transcriptome of *E. coli* O157:H7. Over-expression of *gadE* in non-pathogenic *E. coli* was shown to affect the expression of ~40 genes, including GAD genes (54). Most of these genes, however, differed from those identified following inactivation

of *gadE* in O157:H7, suggesting that apart from its effect on the GAD system, GadE has additional regulatory functions in *E. coli* O157:H7.

Expression of *gadA*, *gadB* and *gadC* were not completely abolished in *E. coli* O157:H7 Δ*gadE* similar to *E. coli* K12 where minimal expression of GadAB proteins was observed in Δ*gadE* (84). Expression of GAD system components in the absence of GadE could be induced by the GadX regulator, which has been shown to bind to and activate *gadA* and *gadBC* transcription directly under in vitro conditions, but not during in vivo growth (132, 159, 162). Another interesting observation was that at stationary phase, the magnitude of increase in expression of *gadA* and *gadB* in the wild type compared to Δ*gadE* were different, indicating that the inactivation of *gadE* affects these duplicated genes in distinct ways. This corroborates the recent finding that the sequences of *gadA* and *gadB* are divergent in *E. coli* O157:H7 in contrast to other *E. coli* strains where gene conversion events between *gadA* and *gadB* have led to genetic homogenization (8). In contrast to the wild type, no increase in expression of GAD genes was observed in Δ*gadE* as the cells entered stationary phase demonstrating that GadE is required for the growth phase regulation of GAD genes.

This study demonstrates that *gadE* inactivation has differential effects on the expression of AFI genes in *E. coli* O157:H7. In non-pathogenic *E. coli* strains, *gadE* induces the expression of AFI genes such as *hdeB*, *hdeA, hdeD, gadX* and *yhiF* in addition to *gadA* (54, 91, 126). In this study, expression of *gadA* and *hdeBAD* showed growth phase-dependent down-regulation in Δ*gadE*. However,

gadX and *yhiF* were not differentially expressed in Δ*gadE* indicating that at pH 7.0, loss of *gadE* does not influence the expression of these two genes in *E. coli* O157:H7. These differences between O157:H7 and non pathogenic strains could also be due to the differences in growth conditions, since strains were grown in rich or minimal media at acidic pH in most of the previous studies.

The relationship between AR and virulence of pathogenic *E. coli*, particularly the interactions between the GAD system and LEE, remains poorly defined. Few studies in the past have shown that some of the GAD system regulators negatively affect LEE expression (24, 99, 138, 157). In *E. coli* O157:H7 Sakai, *gadE* inactivation was found to increase the expression of LEE encoded *espB*, *espD* and *tir* genes, but not *ler* (156). Hence, Mellies et al. considered that GadE mediated down-regulation of LEE was independent of Ler and the pathway through which GadE affects LEE was undetermined (95). Moreover, the extent to which GadE inhibits the transcription of LEE genes has not been demonstrated quantitatively before. The data presented here demonstrate that GadE has a global effect on LEE genes: GadE influences the expression of at least 19 LEE encoded genes belonging to all five LEE operons and 2 non-LEE encoded effectors. These data also provide insight about the mechanism underlying GadE-mediated LEE down-regulation. In contrast to the previous study (156), there was a significant increase in the expression of *ler* in Δ*gadE*, which may be due to differences in growth medium used. The previous study used DMEM containing glycerol (156) whereas in this study MOPS minimal medium was used for growing the cells. This discrepancy could also be due to

difference in sensitivity of the assays used (northern blotting vs. QPCR and microarray). The negative correlation between expression of *gadE* and *ler* was marked in the *gadE*-over expressing strain in which *ler* expression was substantially down-regulated. Furthermore, the identification of a putative GAD box sequence upstream of *ler* with 70% identity to the conserved GAD box sequence provides additional evidence of direct regulation, as GadE has been shown to bind to box sequences with as low as 60% identity to the conserved sequence (85). Additionally, inactivation of *ler* in Δ*gadE* led to a marked decrease in LEE expression, confirming that *ler* is essential for the up-regulation of LEE in Δ*gadE*. Taken together, these findings illustrate that GadE indirectly down-regulates LEE expression most likely through down-regulation of Ler. However, additional putative GAD boxes were observed upstream of other LEE genes, *sepZ* and *escC,* also and therefore, it is possible that GadE directly regulates these LEE genes independent of Ler.

Because GadE negatively influences LEE expression, we hypothesized that environmental conditions that induce *gadE* may down-regulate the expression of LEE. Two conditions that lead to induction of *gadE* are entry into stationary phase and acidic pH (39). Stationary phase expression of LEE genes has been described previously (12). Similarly, influence of several environmental factors such as temperature, bicarbonate ion concentration and membrane stress on the expression of LEE has been investigated (2, 100, 157, 164). However, the effect of pH on LEE expression and the factors regulating that effect in EHEC remain largely unknown. Our experiments demonstrated that

exposure to moderately acidic pH strongly induces *gadE* and has a substantial negative effect on LEE expression in *E. coli* O157:H7. This inhibitory effect could be more profound in extreme acidic conditions such as the gastric environment. To determine whether the acidic pH-induced down-regulation of LEE is exclusively regulated by GadE, we measured the expression of LEE genes in the Δ*gadE* following growth at pH 5.0. Except for *ler,* a partial down-regulation at acidic pH was observed in the expression of the LEE genes in Δ*gadE,* indicating that GadE is not the only regulator responsible for the pH induced down-regulation of LEE. This partial down-regulation of LEE is not mediated through Ler, as expression of *ler* was increased in the Δ*gadE* at pH 5.0. To understand this phenomenon further, we analyzed the expression of other AR regulators that could act on LEE, independent of GadE, and found that *gadX* and *evgA* were also strongly induced at acidic pH in both wild type and Δ*gadE*. GadX has been shown to negatively regulate the expression of LEE genes through the plasmid encoded regulator, Per, in EPEC (138). However, the effect of GadX on expression of LEE in EHEC has not been determined. It is possible that GadX regulates LEE through an unknown regulator in EHEC. The decrease in LEE gene expression in an acidic environment in the Δ*gadE* is likely to be mediated by EvgA also. Previously, EvgA has been shown to repress LEE, independent of Ler, by activating *ydeO* and *ydeP* (99). Because YdeO is a positive regulator of *gadE,* the partial decrease in LEE expression in Δ*gadE* may occur through YdeP. Collectively these experiments suggest that GadE, GadX and EvgA may cooperatively repress the expression of LEE genes at acidic pH and that GadE is

the sole regulator responsible for the changes in expression of *ler* in the acidic conditions used in this study.

The AR of Δ*gadE* also was characterized by assessing survival at pH 2.0 and 2.5 in the minimal AR mechanism environments and in the complex acidic conditions of the MSS. The Δ*gadE* failed to survive the acid challenge at pH 2.0 + glutamate, indicating lack of a functional GAD system. Inactivation of *gadE* in *E. coli* O157:H7 negatively impacted the protective ability of the ARG and OXI AR systems. The survival of Δ*gadE* was tested in the MSS, which evaluates the ability of bacterial strains to survive in a gastric environment after ingestion of food (64). The GadE central regulator, and thus a functional GAD system, is a critical component for survival, as inactivation of *gadE* abrogated the ability of *E. coli* O157:H7 to survive in the MSS. Hence, GadE most likely plays a protective role during the passage of O157 through the gastric environment. This assumption is supported by a previous study by Price et al., (119) which demonstrated that *gadC* is required for the survival of *E. coli* O157:H7 in calves.

In summary, this study shows that GadE is an important regulator that modulates the expression of AR and virulence genes in *E. coli* O157:H7 in response to environmental conditions similar to those that are found in various food matrices and the human gastrointestinal tract. GadE has acquired additional functions in *E. coli* O157:H7 and it acts as a link between AR and virulence: it activates the GAD system of AR and at the same time down-regulates the expression of LEE genes, which are important for the adhesion of the organism to intestinal mucosa and development of AE lesions.

Consequently, we propose that during passage through the human stomach GadE protects *E. coli* O157:H7 by inducing the GAD system and aids in energy conservation by inhibiting the unnecessary expression of the LEE genes. As the organism reaches the intestine, environmental changes including alkaline pH and high $NaHCO_3$ concentration induce the LEE regulator, Ler, which negatively regulates expression of *gadE* (1) leading to inhibition of the GAD system.

ACKNOWLEDGMENTS

This work was supported by Food Safety NRI #2005-35201-16362 from the United States Department of Agriculture, and in part by funds from the NIAID, NIH, DHHS, under the Food and Waterborne Diseases Integrated Research Network (NIH Research Contract N01-AI-30058).

The authors wish to thank Vanessa Sperandio, University of Texas South Western Medical center, for helpful suggestions for creating genetic complements and Galeb Abu-Ali, James T. Riordan and Shannon Manning for critical review of the manuscript.

CHAPTER 3

Effect of acidic pH on the expression of virulence factors of *Escherichia coli* O157:H7: role of glutamate decarboxylase system regulators

SUMMARY

Effect of gastric acidity on the expression of virulence factors remains less investigated in an important enteric pathogen, *Escherichia coli* O157:H7. We hypothesized that exposure to extreme acidity downregulates the expression of virulence factors in *E. coli* O157:H7. Stationary phase cells of *E. coli* O157:H7 were exposed to pH 2.0 and pH 7.0 for 10 min to analyze the expression of virulence and acid resistance genes. Supporting the hypothesis, exposure to pH 2.0 resulted in a marked decrease in the expression of the major virulence factors of *E. coli* O157:H7 such as Shiga toxins, locus of enterocyte effacement (LEE) and pO157 plasmid encoded genes. On the other hand, acid resistance genes such as *gadA*, *gadB*, and *gadC* showed increased expression at pH 2.0. Furthermore, we investigated whether two acid resistance regulators that negatively regulate LEE expression in enteropathogenic *E. coli* (EPEC), GadX and EvgA, have a similar effect on LEE expression in *E. coli* O157:H7. Contrary to EPEC, inactivation of *gadX* had only a minimal effect on the expression of LEE genes in *E. coli* O157:H7 whereas inactivation of *evgA* resulted in a marked increase in expression of the LEE genes, *tir*, *espD*, *eae*, and *sepZ* similar to EPEC.

E. coli O157:H7 is the most prevalent enterohemorrhagic *E. coli* (EHEC) serotype in the United States (67). It is a food borne zoonotic pathogen and causes hemorrhagic colitis and hemolytic uremic syndrome (HUS) in human beings (153). HUS is a thrombotic disorder that causes kidney damage in 15% of the children infected with *E. coli* O157:H7 (153). Before colonizing the lower intestinal tract of humans, *E. coli* O157:H7 has to pass through the stomach, which presents one of the first host defense barriers. Gastric acidity (pH 1.5 – 2.5) is lethal to most bacteria including many enteric pathogens (112). *E. coli* strains are extremely acid resistant and can survive gastric acidity efficiently (39). At least four mechanisms contribute to this acid resistance of *E. coli*, among which glutamate decarboxylase (GAD) system is the most efficient (39).

The effect of acidity on the expression of virulence factors of *E. coli* O157:H7 is largely unknown. Major virulence factors of *E. coli* O157:H7 include one or more Shiga toxins (Stx), a pathogenicity island namely locus of enterocyte effacement (LEE) and a plasmid, pO157 (170). Several environmental factors such as growth phase, temperature, presence of iron, osmolarity, and presence of antibiotics have been shown to affect the expression of these virulence factors (12, 20, 95). However, the effect of acidity on the expression of Stx, LEE and pO157 genes remains less clear. Therefore, a study on the expression of virulence factors at acidic pH is critical as it could provide important insights about how *E. coli* O157:H7 synchronizes the expression of acid resistance genes and virulence genes when exposed to extreme acidic environments such as the human stomach. In this context, we hypothesized that upon exposure to extreme

acidity such as pH 2.0, *E. coli* O157:H7 downregulates its virulence factors and upregulates the acid resistance genes in order to facilitate survival and conserve energy.

In a previous study, we determined the effect of acidity on LEE expression by growing the *E. coli* O157:H7 cells at pH 5.0 and demonstrated a marked decrease in expression of *ler*, *espD*, and *sepZ*, at pH 5.0 (65). However, the expression pattern of genes observed at pH 5.0, although interesting, may not be representative of the expression at gastric pH as pH 5.0 represents only moderate acidic conditions. Moreover, the change in expression of other major virulence factors such as Stx and pO157-encoded genes at acidic pH was not investigated. Therefore, to determine the effect of gastric pH on the expression of virulence and acid resistance genes, we exposed stationary phase (OD_{600} ~ 1.5) cells of *E. coli* O157:H7 Sakai strain to EG minimal medium containing glutamate (65, 83) at pH 2.0 and pH 7.0 for 10 min and compared the expression of selected virulence genes. The virulence genes tested were LEE genes, *ler*, *tir*, *sepZ* (*espZ*), and *espD* that are involved in causing attaching and effacing lesions in the intestine; Stx genes, *stx1* and *stx2*, which encode two variants of Shiga toxins; pO157 genes, *ehxA*, which encodes enterohemolysin, and *etpC*, a type II secretion system gene involved in adhesion to the epithelial cells. Additionally, to determine the effect of pH 2.0 on acid resistance genes, the expression of three GAD system genes *gadA*, *gadB* and *gadC* was analyzed. For all the experiments described in this report, RNA was extracted using modified hot phenol extraction method as described previously (9). One microgram of

RNA was converted into cDNA, which was then used for quantitative real time PCR (qRT-PCR) as described before (65). The fold change in expression of genes was calculated by Pfaffl's method (113).

Exposure to pH 2.0 altered the expression of all the genes tested. There was a 2 – 5-fold upregulation of the three components of GAD system, *gadA*, *gadB* and *gadC*. This is an expected result as the GAD system has been shown to provide maximum protection to the cells at extreme acidity (39, 75). In contrast, all of the virulence genes tested showed decrease in expression of different magnitudes at pH 2.0 with some of them showing a dramatic decrease in expression. *stx1* had a 28.2-fold decrease in expression whereas *stx2* was down regulated by only 2.3-fold upon exposure to pH 2.0. Similarly, the LEE genes tested demonstrated decreased expression at pH 2.0, among which the highest repression of 74.1- and 60.9-fold were observed for *tir* and *ler*, respectively. *espD* had a 23.6-fold decrease in expression whereas *sepZ* showed only a 2-fold decrease in expression at pH 2.0. The two pO157 plasmid-encoded genes *ehxA* and *etpC* also showed decrease in expression of 16.2- and 9.6-fold, respectively (Table 3.1).

In our previous study (65), the central activator of the GAD system, GadE (39), was involved in the decrease in expression of the LEE genes at acidic pH in *E. coli* O157:H7. However, inactivation of *gadE* only partially relieved the acid-induced repression of LEE and therefore, in this study, we hypothesized that additional regulators such as GadX and EvgA are also involved in this

TABLE 3.1

Alteration in expression of acid resistance and virulence genes upon exposure to extreme acidity

Gene	Fold change (pH 2.0/pH 7.0)*
gadA	2.6±0.8
gadB	2.3±0.7
gadC	5.2±1.4
sepZ	-2±0.2
tir	-74.1±14.5
espD	-23.6±3.3
ler	-60.9±29.2
hlyA	-16.2±7.8
etpC	-9.6±0.9
stx1	-28.2±16.3
stx2b	-2.3±0.3

* Positive values indicate increased expression at pH 2.0 and negative values indicate decreased expression at pH 2.0.

phenomenon as GadX and EvgA have been previously shown to negatively regulate LEE expression in EPEC (99, 139).

To determine whether GadX and EvgA regulate the expression of LEE in EHEC, Δ*gadX* and Δ*evgA E. coli* O157:H7 Sakai strains were constructed through a one step inactivation method (32, 98) and the LEE expression was analyzed by qRT-PCR. Wild type (WT) and Δ*gadX* strains were grown in previously described *gadX*-inducing conditions (39, 139) to analyze the expression of LEE genes. This included growth in morpholino propane sulfonic acid (MOPS) minimal medium at pH 7.4 and Dulbecco's modified Eagle's medium (DMEM) at pH 5.5 (0.2% glucose, 3.7% NaHCO3, pH adjusted with morpholino ethane sulfonic acid (MES)) upto an OD_{600} of 0.5. Expression analysis was not conducted at pH 2.0 as Δ*gadX* is not viable at this pH.

Inactivation of *gadX* resulted in a minimal increase in expression of some of the LEE genes at neutral pH (Table 3.2). Interestingly, this effect was not observed at pH 5.5 (Table 3.2). This result is in contrast with that of EPEC, where inactivation of *gadX* had a significant effect on LEE expression at pH 5.5 but not at neutral pH. GadX negatively regulates LEE expression in EPEC by downregulating a plasmid encoded positive regulator of LEE, PerA (139). No PerA homolog has been identified in EHEC (61) and that could be the reason for the observed differences in GadX-mediated LEE regulation in *E. coli* O157:H7.

Subsequently, the expression of LEE genes in Δ*evgA* was analyzed in previously described *evgA*-inducing conditions (39). The cells were grown to exponential phase (OD_{600} = 0.5) in minimal medium (E minimal medium with

TABLE 3.2

Effect of *gadX* inactivation on the expression of LEE genes

Gene	Δ*gadX*/WT (pH 7.4)	Δ*gadX*/WT (pH 5.5)
tir	1.1±0.2	0.5±0.1
espD	1.8±0.1	0.8±0.1
eae	1.5±0.2	0.8±0.2
sepZ	1.5±0.3	0.8±0.2

0.2% glucose and 12mM $MgSO_4$) at pH 7.0 and pH 5.0. There was a marked increase in expression of LEE genes in the Δ*evgA* (Table 3.3) at neutral pH indicating that a negative interaction between EvgA and LEE, similar to that in EPEC, exists in *E. coli* O157:H7 as well. Interestingly, this effect was not as pronounced at pH 5.0 as at neutral pH (Table 3.3). As in EPEC (99), this negative regulation of LEE by EvgA could be mediated through YdeO and GadE regulators.

The ability to orchestrate gene expression in response to environmental conditions is an important requirement for being a successful pathogen. Expression of virulence factors in inappropriate environments, particularly where they are not needed, can negatively affect energy conservation and markedly impair survival of the pathogen in the host. It has been shown in *Bordetella bronchiseptica* that ectopic expression of flagellar regulon interferes with the tracheal colonization by the pathogen (4). During passage through human stomach, the main objective of *E. coli* O157:H7 is its survival and therefore, it is expected that genes that provide protection at extreme acidity are positively selected for upregulation and at the same time, expression virulence factors that are not required in that environment is suppressed. The results from this study support this argument; upon exposure to gastric acidity *E. coli* O157:H7 upregulated acid resistance genes such as GAD genes and downregulated major virulence factors such as Stx, LEE and pO157 genes. This may be beneficial to the organism as decreased LEE expression prevents attachement of the bacteria to the gastric mucosa and allows for faster transit through stomach.

TABLE 3.3

Effect of *evgA* inactivation on the expression of LEE genes

Gene	Δ*evgA*/WT (pH 7.0)	Δ*evgA*/WT (pH 5.0)
tir	3.7±0.5	1.4±0.5
espD	4.4±0.5	1.2±0.3
eae	3.8±0.6	1.0±0.1
sepZ	5.3±1.5	1.2±0.4

Additionally, this pattern of gene expression may help in maximum energy conservation by the organism. We hypothesize that GAD system regulators such as GadE and EvgA are involved in the decrease in LEE expression at extreme acidity. It is possible that the EvgAS two-component system senses the acidity in the environment and upregulates the expression of GadE, which in turn induces GAD expression and suppresses LEE expression. We do not know the factors that are involved in the repression of Stx and pO157 genes at extreme acidity at this time because none of the GAD system regulators investigated in this study have been shown to regulate their expression. Further work is needed in this aspect to identify the genes controlling this response and the mechanism of downregulation.

This work was supported by Food Safety NRI #2005-35201-16362 from the United States Department of Agriculture, and in part by funds from the NIAID, NIH, DHHS, under the Food and Waterborne Diseases Integrated Research Network (NIH Research Contract N01-AI-30058).

CHAPTER 4

Differential Expression of Virulence and Stress Fitness Genes between Clinical and Bovine-biased Genotypes of *Escherichia coli* O157:H7

This chapter is submitted to Applied and Environmental Microbiology

SUMMARY

Escherichia coli O157:H7 strains can be classified into different genotypes based on the presence of specific Shiga toxin-encoding bacteriophage insertion sites. Certain O157:H7 genotypes predominate among human clinical cases (clinical genotypes), while others are more frequently found in bovines (bovine-biased genotypes). To determine whether inherent differences in gene expression explains the variation in infectivity of these genotypes, we compared the expression patterns of clinical genotype 1 strains with those of bovine-biased genotype 5 strains using microarrays. Important O157:H7 virulence factors including locus of enterocyte effacement genes, the enterohemolysin, and several pO157 genes, showed increased expression in the clinical versus bovine-biased genotype. In contrast, genes essential for acid resistance (e.g., *gadA*, *gadB*, and *gadC*) and stress fitness were upregulated in bovine-biased genotype 5 strains. Increased expression of acid resistance genes was confirmed functionally using a model stomach assay, in which strains of bovine-biased genotype 5 had a 2-fold higher survival rate than strains of clinical genotype 1. Overall, these results suggest that the increased prevalence of O157:H7 illness caused by clinical genotype 1 strains is due in part to the overexpression of key virulence genes. The bovine-biased genotype 5 strains, however, are more resistant to adverse environmental conditions, a characteristic that likely facilitates asymptomatic O157:H7 colonization of bovines.

INTRODUCTION

Escherichia coli O157:H7, a food-borne zoonotic pathogen that causes hemorrhagic colitis and hemolytic uremic syndrome (HUS) in humans, is the most prevalent type of enterohemorrhagic *E. coli* (EHEC) in the United States (67, 153). Cattle are the primary reservoir of *E. coli* O157:H7 and the fecal shedding rate on cattle farms can be up to 100% (45). While colonized cattle do not exhibit clinical disease (127), it has been reported that only 10-100 cells of *E. coli* O157:H7 are sufficient to induce overt disease in humans (163). Although it has been suggested that bovine-derived *E. coli* O157:H7 strains vary in their ability to cause human disease (14), the basis behind this variation is not known.

E. coli O157:H7 possesses unique virulence properties that facilitate disease development including Shiga toxins (Stx), the locus of enterocyte effacement (LEE) pathogenicity island, and the pO157 virulence plasmid (171). The LEE encodes a type 3 secretion system (T3SS) that mediates the formation of attaching and effacing lesions (94), while the pO157 plasmid encodes several putative virulence factors such as an enterohemolysin (EhxA or EHEC-HlyA) (78) and a type 2 secretion system (T2SS) (133). Both the LEE and pO157 have shown to be critical for disease pathogenesis (66, 81). Shiga toxins, which are the cytotoxins responsible for renal damage in HUS (66), are encoded by genes located on lysogenic lambdoid phages that are inserted into the O157:H7 chromosome at specific locations (136). A prior study of 80 bovine isolates and 282 clinical isolates from humans with O157-associated disease demonstrated that the distribution of Stx insertion sites varies between isolate types (14).

Furthermore, these isolates were classified into different genotypes based on the insertion sites of Stx-encoding bacteriophages and genotypes 1 – 3, although isolated from cattle also, were predominant in clinical isolates and were considered as clinical genotypes. On the other hand, genotypes such as 5 and 7 were overrepresented among the cattle isolates and therefore, considered as bovine-biased genotypes (14). Similarly, octamer-based genome scanning of bovine and clinical isolates of *E. coli* O157:H7 identified two genetically distinct lineages, of which lineage I was isolated mostly from humans and lineage II, mostly from bovines (72). Comparing the presence of virulence genes between *E. coli* O157:H7 isolates from various sources using DNA microarrays also has revealed that O157 isolates from beef cattle and humans are genetically distinct (79).

In addition to intrinsic differences, it is possible that there are differences in the expression of important virulence genes as well as variation in the degree of resistance to adverse environmental conditions between clinical and bovine-biased genotypes. To investigate this hypothesis, the exponential phase transcriptomes of four clinical genotype 1 strains were compared to the transcriptomes of four bovine-biased genotype 5 strains using microarrays. All the strains used in this study were from bovine sources and they belonged to either genotype 1 or 5. Therefore the Genotype 1 strains used in this study are in fact bovine-derived clinical genotype strains. The goal of this study was to identify specific genes that are differentially expressed between the two genotypes to better understand why genotype 1 strains cause more clinical

disease than genotype 5 strains. The identification of genes that are upregulated in clinical versus bovine-biased genotypes is important for detecting and controlling those strains that are more likely to cause *E. coli* O157:H7 infections in humans.

MATERIALS AND METHODS

Bacterial strains. The eight bacterial strains used in this study were selected based on the Stx-encoding bacteriophage insertion site genotypes determined in a prior study (14). Strains representing genotypes 1 (clinical genotype) and 5 (bovine-biased genotype) were selected among 80 bovine strains originally isolated between 1991 and 2004 as described (14). Although the clinical genotype 1 strains in this study were bovine-derived, their genotype was identical to those genotype 1 strains isolated in a prior study from humans with O157 infections (14). Four strains of each genotype were included in the microarray analyses. A previously described (88) *stx2/stx2c* RFLP demonstrated that all the genotype 1 strains used in this study harbored *stx2* alone, whereas the genotype 5 strains contained only *stx2c*.

Growth conditions. Each strain was stored at -70°C in LB broth containing 10% glycerol, was inoculated into 10 ml LB broth, and grown to an OD_{600} of ~0.1 to recover cells. Cells were grown twice to stationary phase in MOPS-buffered minimal medium (pH 7.4) before transferring at a 1:30 dilution into 100 ml Dulbecco's Modified Eagle Medium (DMEM) (0.45% glucose) for both RNA isolation and the model stomach assay. To minimize the confounding effect of acidic pH that develops in stationary phase of growth in un-buffered medium, the DMEM was buffered with MOPS to pH 7.4.

Microarray design. To compare global gene expression profiles between genotypes 1 and 5, microarrays were hybridized in a double loop design, thereby allowing strains from one genotype to be directly compared to strains from the

other genotype (Fig. 4.1). The four strains from each genotype represent biological replicates and therefore, significant differences in gene expression are representative of the two genotypes.

RNA isolation and cDNA labeling. For RNA isolation, the strains were grown to exponential phase (~2.25 h, OD_{600} ~0.5) in DMEM and RNA extractions were performed using a modified version of the previously described hot-phenol method (15). Briefly, 5 ml of the culture was mixed with 1/10V of 10% phenol:ethanol buffer to stabilize the RNA, and centrifuged at 4° C (4300 × g) for 30 min to pellet cells. The supernatant was decanted and cell pellets were suspended in 5 ml of buffer (2 mM EDTA, 20 mM NaOAc, pH 5.2) before RNA extraction with hot-phenol. Reverse transcription reactions and the coupling of cDNA with Cy3 or Cy5 dyes were conducted as described elsewhere (12).

cDNA hybridizations. Hybridizations were performed according to the double loop microarray design (Fig. 1). As described (12), the cDNAs were hybridized onto microarray slides printed with 6,088 ORFs representing *E. coli* genome strains K12 (17), EDL933 (110) and Sakai (46);110 ORFs from the pO157 plasmid were included. Arrays were scanned with an Axon 4000b scanner (Molecular Devices, Sunnyvale, CA) followed by image analysis using GenePix 6.0 (Molecular Devices).

Data analysis. Microarray data were processed as previously described (65) and fitted to a mixed ANOVA model (30). The linear model tested was Y (intensity) = array + dye + strain (clinical or bovine-biased) + sample (biological replicate) + error. Significant differences in expression were determined using the

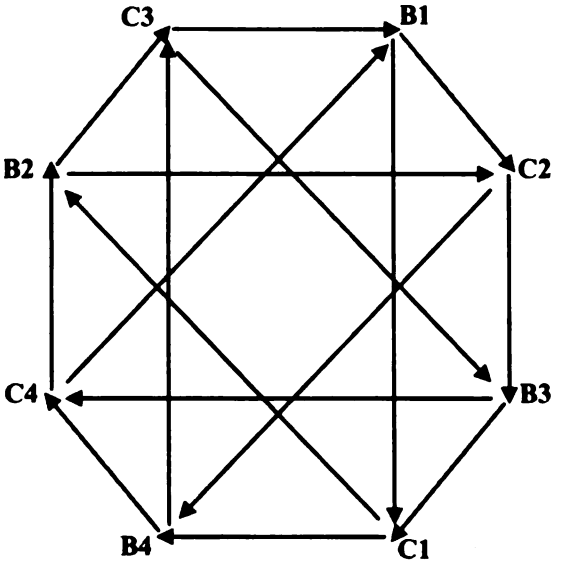


Figure 4.1. Double loop design for the microarray experiment. B1-B4 represent four different strains of the bovine-biased genotype 5, whereas C1-C4 represent four different strains of the clinical genotype 1. Each arrow indicates a hybridization with the arrow head representing Cy3 and tail Cy5 dyes.

Fs test in MAANOVA with 500 random permutations to estimate the p-values. This test uses a shrinkage estimator for gene-specific variance components that makes no assumption about the variance across genes (31). In addition, the q-value package in R was used to determine the false discovery rate (FDR) (146). Additionally, significance analysis of microarrays (SAM) was used to analyze data with a FDR of 0.05.

Overrepresentation of gene sets with a common biological function in the two genotypes was determined using the Gene Set Enrichment Analysis (GSEA) Preranked analysis program (GSEA v2.0; Broad Institute, Massachusetts Institute of Technology) (148). The gene sets were designated based on the annotation for the Sakai genome (46) available through the J. Craig Venter Institute (http://cmr.jcvi.org/tigr-scripts/CMR/GenomePage.cgi?org=ntec03). Additionally, genes for the LEE and the AFI-GAD, which represents the glutamate decarboxylase (GAD) system and acid fitness island (AFI), also were included in the analysis.

Quantitative real-time PCR (qRT-PCR). Select genes that had significantly different levels of expression between genotypes in the microarray analysis were confirmed by qRT-PCR. Taqman assays (11) were used to quantify the expression of *gadA*, *gadB*, and *ler*, with *mdh* as a reference for normalization. For all other genes, SYBR green was used as described elsewhere (65); methods for cDNA synthesis and qRT-PCR also were described previously (65). The expression level of the 16S rRNA gene was used for normalization of data and the relative expression levels were quantified using

modified Livak method (134). The results presented are averages from four biological replicates ± standard error of mean (SEM).

Single Nucleotide Polymorphism (SNP) Genotyping. Genomic DNA was extracted with the Purgene DNA extraction kit (Gentra systems, Minneapolis, MN) for use with the GenomeLab SNPstream™ system (Beckman Coulter, Fullerton, CA). SNP genotyping via the SNPstream™ was performed using a modified version of a previously described protocol (88) according to the manufacturer's instructions. Briefly, PCR was conducted using four panels of 48-plex primers targeting 192 distinct SNP loci identified previously via the comparison of three O157:H7 genomes (88). The primers, which differed from the original protocol, were designed using the Autoprimer program (175). After cleaning, the PCR products were subjected to single base primer extension reactions that add a labeled nucleotide to the SNP site followed by hybridization onto a 384-well SNP microarray plate. Detection and processing were performed via the SNPstream™ Imager (version 2.3; GenomeLab). A total of 52 of the 192 SNPs were found to be informative; these SNPs were concatenated in MEGA4 (151) to construct a neighbor-joining tree (130) for examining the phylogenetic relationships between the eight strains. SNP data from reference strains representing each of the nine O157:H7 clades (88) were included in the analysis.

Model stomach assay. The model stomach system (MSS) (64) was prepared as described previously (11). Gerber Turkey Rice Dinner© baby food (30 g) was mixed with 120 ml of synthetic gastric fluid (pH 1.70), yielding a final pH of 2.5. Strains, grown to an OD_{600} ~2.5 in DMEM, were inoculated into the

MSS at the rate of 10^6 cells/ml. Contents of the MSS were stomached for 30 sec, sampled, diluted, and plated onto LB agar plates every 30 min for 1.5 h to enumerate viable cells. CFU/ml from duplicate plates were averaged and converted to log10 CFU/ml. Survival rates were calculated as the log decrease in viable cell count per 30 min, and the average from two experimental replicates were reported.

Microarray data accession. Microarray data are available at NCBI GEO (http://www.ncbi.nlm.nih.gov/geo), accession number GSE15783.

RESULTS

Differentially expressed genes between clinical and bovine-biased genotypes of *E. coli* O157:H7. The Fs test identified significant differential expression of 191 genes between the two genotypes, of which 71 were upregulated in the clinical genotype and 120 were upregulated in the bovine-biased genotype (FDR<0.1, fold change ≥ 1.5) (Table 4.S1 in appendix). Additionally, SAM identified more genes to be significantly differentially expressed between the two genotypes; 154 were upregulated in the clinical and 238 were upregulated in the bovine-biased genotype (FDR<0.05, fold change >1.5) (Table 4.S2 in appendix). One hundred and sixteen genes were found to be differentially expressed by both MAANOVA and SAM. Differentially expressed genes included those involved in virulence, response to stress, acid resistance and metabolism. Overall, important O157:H7 virulence factor genes were upregulated in the clinical genotype (Table 4.1), whereas genes related to acid resistance and stress fitness were upregulated in the bovine-biased genotype (Table 4.2). GSEA also identified enrichment of 8 gene sets in the clinical genotype and 6 gene sets in the bovine-biased genotype (Table 4.3), thereby providing additional support for the MAANOVA and SAM results.

The LEE genes. There was an overall increase in expression of the LEE genes in clinical genotype 1 relative to bovine-biased genotype 5 strains; expression was significantly different in12 genes (Table 4.1). For example, secreted proteins encoded by *espF* and *espG*, T3SS proteins encoded by *escF*, *sepQ*, *escT*, and *escR*, and the *cesD* chaperone were upregulated in genotype 1

TABLE 4.1

Virulence-associated genes upregulated in clinical genotype 1 relative to bovine-biased genotype 5

ECs no.[a]	Gene	Function	Expression ratio (C/B)[b]	Test[c]
LEE genes				
ECs4550	*espF*	Effector protein	2.0	S
ECs4552	*escF*	T3SS EscF protein	2.1	S
ECs4565	*sepQ*	T3SS structural protein	1.6	M
ECs4566	*orf16*	Secretion of translocators	2.2	S
ECs4570	*orf12*	T3SS	2.1	S
ECs4574	*sepD*	T3SS SepD protein	2.0	S
ECs4576	*cesD*	T3SS chaperone	1.8	M
ECs4579	*rorf3*	T3SS	1.9	S
ECs4581	*escT*	T3SS structural protein	2.0	S
ECs4583	*escR*	T3SS structural protein	1.9	S
ECs4590	*espG*	Effector protein	1.9	S
ECs4591	*rorf1*	Unknown	1.6	M
Genes encoded by pO157 plasmid				
pO157p02	*etpC*	T2SS	2.8	M
pO157p05	*etpF*	T2SS	1.9	M+S
pO157p07	*etpH*	T2SS	1.9	M
pO157p09	*etpJ*	T2SS	2.0	M+S
pO157p10	*etpK*	T2SS	2.2	M+S
pO157p11	*etpL*	T2SS	1.8	M
pO157p12	*etpM*	T2SS	1.9	M+S
pO157p14	*etpO*	T2SS	1.7	M
pO157p18	*ehxA*	enterohemolysin	1.7	M
pO157p24	*repFIB*	replication protein	1.7	M
pO157p58	*toxB*	toxin B	1.6	M
pO157p79		unknown	2.4	M+S
pO157p80		unknown	2.1	S
pO157p81		unknown	2.1	M+S

[a] Locus ID for *E. coli* O157:H7 Sakai strain (Genbank # BA000007)

[b] Expression ratio between two genotypes; C=clinical genotype, B=bovine-biased genotype

[c] Test that identified a gene as statistically significant; M=MAANOVA, S=SAM

TABLE 4.2

Genes upregulated in bovine-biased genotype 5

ECs no. [a]	Gene	Function	(B/C) [b]	Test [c]
GAD and AFI genes				
ECs2097	*gadC*	GABA-glutamate antiporter	2.5	M+S
ECs2098	*gadB*	Glutamate decarboxylase isozyme	4.4	M+S
ECs4377	*slp*	Outer membrane protein	2.5	M+S
ECs4389	*hdeB*	Periplasmic chaperone	3.4	M+S
ECs4390	*hdeA*	Protection from organic acid metabolites	5.6	M+S
ECs4391	*hdeD*	Acid resistance at high cell density	2.7	M+S
ECs4392	*gadE*	Central activator of the GAD system	2.6	S
ECs4394	*yhiV*	Multi drug efflux pump protein	2.2	M+S
ECs4395	*gadW*	ARAC-type GAD system regulator	1.8	M+S
ECs4397	*gadA*	Glutamate decarboxylase isozyme	1.9	M+S
Stress fitness-associated genes				
ECs0890	*dps*	DNA protection during starvation	3.6	M+S
ECs0966	*cspD*	Cold shock protein	2.5	M+S
ECs1722	*chaB*	Cation transport regulator	2.4	M+S
ECs4871	*katG*	Catalase-peroxidase	1.7	M+S
ECs2086	*osmC*	Osmotically inducible protein	2.1	M+S
ECs5334	*osmY*	Osmotically inducible protein	3.0	M+S
ECs4367	*uspA*	Universal stress protein	2.6	S
ECs0968	*clpA*	Degradation of abnormal proteins	2.0	S
ECs1723	*chaC*	Cation transport regulator	2.2	S
RpoN regulated (involved in nitrogen metabolism)				
ECs0169	*glnD*	uridylyltransferase acts on regulator of glnA	1.8	S
ECs0504	*glnK*	nitrogen regulatory protein P-II 2	11.6	S
ECs0505	*amtB*	probable ammonium transporter	14.9	S

Table 4.2 continued...

ECs0692	*gltK*	glutamate/aspartate transport system permease	1.9	S
ECs0693	*gltJ*	glutamate/aspartate transport system permease	2.2	S
ECs0889	*glnH*	permease of periplasmic glutamine-binding protein	2.1	S
ECs2784	*nac*	nitrogen assimilation control protein	12.7	S
ECs3194	*argT*	periplasmic arginine binding protein	1.5	M
ECs4091	*gltB*	glutamate synthase, large subunit	1.9	S
ECs4790	*glnG*	nitrogen regulator I	2.6	S
Nac regulated				
ECs1743	*oppA*	oligopeptide transport	2.1	M+S
ECs1744	*oppB*	oligopeptide transport	1.6	M+S
ECs1746	*oppD*	oligopeptide transport	1.8	M+S
ECs1747	*oppF*	oligopeptide transport	1.7	M+S
ECs3522	*gabD*	succinate-semialdehyde dehydrogenase	3.7	S
ECs3523	*gabT*	4-aminobutyrate aminotransferase activity	3.2	S
ECs4424	*dppA*	dipeptide transport protein	1.9	S
RpoN-regulated (not involved in nitrogen metabolism)				
ECs3582	*hypA*	pleiotrophic effects on hydrogenase isozymes	3.0	S
ECs3583	*hypB*	hydrogenase isoenzyme HypB	2.6	S
ECs5061	*fdhF*	subunit of formate dehydrogenase H	2.1	S

[a] Locus ID for *E. coli* O157:H7 Sakai strain (Genbank # BA000007)

[b] Expression ratio between two genotypes; B=bovine-biased genotype, C=clinical genotype

[c] Test that identified a gene as statistically significant; M=MAANOVA, S=SAM

TABLE 4.3

Enrichment of gene sets in clinical and bovine-biased genotypes

Gene set	NES[a]	FDR-q value[b]
Enriched in clinical genotype 1		
LEE	2.58	0.000
Folic acid biosynthesis	1.88	0.008
tRNA rRNA base modification	1.87	0.006
Pyrimidine ribonucleotide biosynthesis	1.82	0.013
RNA processing	1.74	0.026
Nucleotide and nucleoside interconversions	1.69	0.037
Toxin production and resistance	1.68	0.036
Ribosomal protein synthesis and modification	1.65	0.040
Enriched in bovine-biased genotype 5		
AFI-GAD	-2.09	0.000
Adaptations to atypical conditions	-1.94	0.004
Glutamate family	-1.90	0.004
Pyruvate family	-1.89	0.004
Glycolysis/gluconeogenesis	-1.89	0.003
Fermentation	-1.78	0.010

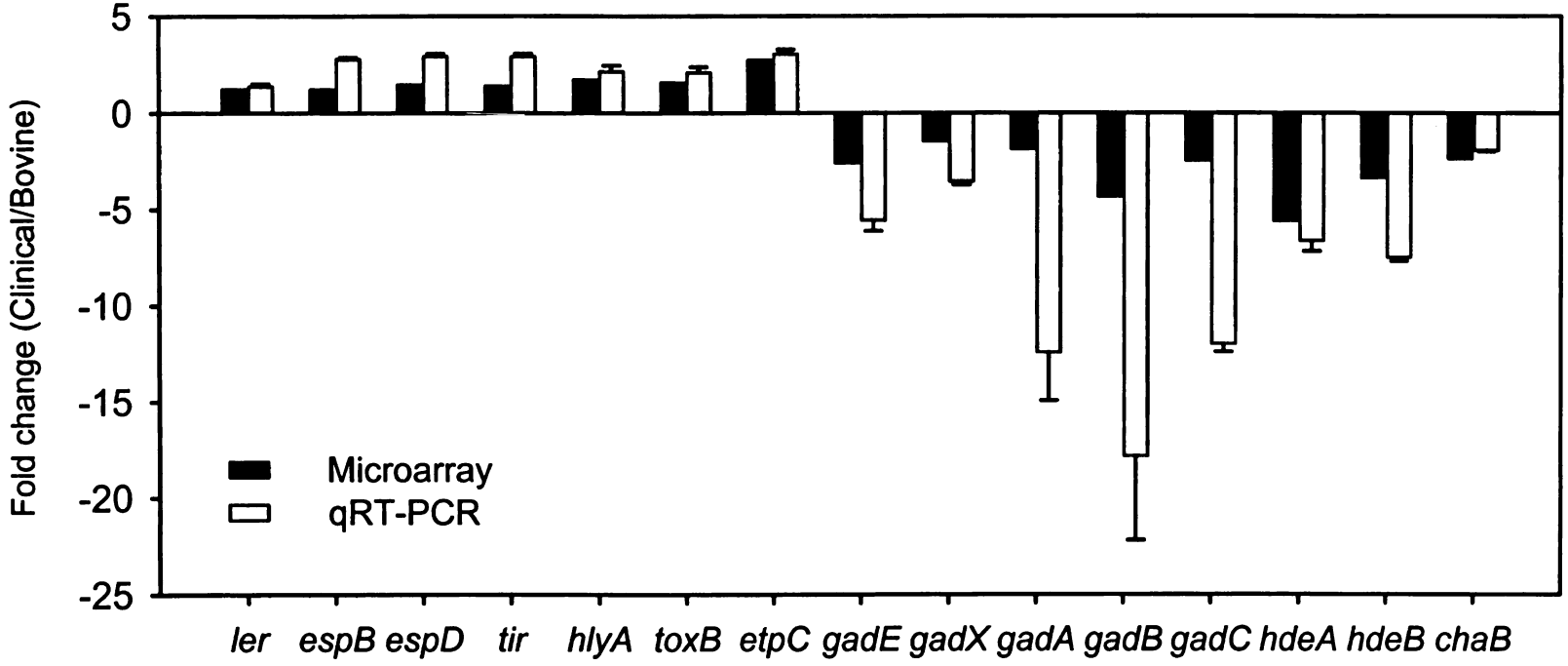
[a] NES= normalized enrichment score indicating the degree to which a gene set is overrepresented in the top or bottom of the ranked list of genes.

[b] FDR-q value ≤ 0.05, indicates false discovery rate.

strains (Table 4.1). Although the remaining 29 LEE genes were not significantly different between genotypes, 27 were upregulated in genotype 1 relative to genotype 5. Members of all the five LEE operons showed change in expression in the same direction. The insignificant result is possibly due to inter-strain variation within the genotypes. To confirm expression differences, qRT-PCR was used to examine the expression of four important LEE genes including *ler*, *espB*, *espD* and *tir* that were not significantly different by microarrays (Fig. 4.2). More than a two-fold increase in expression was observed for *espB*, *espD* and *tir* by qRT-PCR in the clinical genotype (Fig. 4.2), a level that was similar to the microarray data for the 12 significant genes. Expression of *ler* was slightly lower than the other genes, though it still exhibited a 1.4-fold increase in clinical strains (Fig. 4.2). Additionally, GSEA confirmed the enrichment of the entire set of 41 LEE genes in the clinical genotype (Table 4,3).

pO157 plasmid encoded genes. The pO157 plasmid encodes a number of virulence associated genes in *E. coli* O157:H7 strains. Fourteen of these genes, which includes *ehxA* (EHEC-*hlyA*; enterohemolysin), *toxB* (toxin B), and eight of the thirteen genes that encode the T2SS were significantly upregulated in the clinical genotype 1 (Table 4.1). The T2SS *etp* cluster (133) showed a 1.8- to 2.8-fold increase in expression in the clinical genotype, which was confirmed by qRT-PCR (Fig. 4.2). The *ehxA* and *toxB* also were confirmed to have a 2.2- and 2.1-fold increase in expression by qRT-PCR (Fig. 4.2). Expression of some virulence genes encoded by pO157 such as *ehxCBD* and *stcE*, were not significantly different between the two genotypes.

Figure 4.2. qRT-PCR validation of microarray data. The expression ratio between clinical and bovine-biased genotypes as calculated by microarrays and qRT-PCR are given. Results shown are average fold change in expression with standard error of mean (SEM) from four biological replicates (strains) per genotype.


Fold change (Clinical/Bovine)
5
0
-5
-10
-15
-20
-25
Microarray
qRT-PCR
ler
espB
espD
tir
hlyA
toxB
etpC
gadE
gadX
gadA
gadB
gadC
hdeA
hdeB
chaB

Acid resistance and stress fitness-associated genes. Numerous genes that are essential for acid resistance in *E. coli* were significantly upregulated in the bovine-biased genotype 5 strains. This included genes that encode all three components of the GAD system, *gadA*, *gadB*, and *gadC* (39). In addition, the twelve AFI genes (91) had increased expression in the bovine-biased genotype, with eight of the 12 having significantly different levels (Table 4.2). The increased expression of GAD system genes was confirmed by qRT-PCR, which showed a more than 10-fold increase relative to clinical genotype 1 (Fig. 4.2). Similarly, there was a 3.6-, 5.6-, 6.6- and 7.5-fold increase in expression of *gadX*, *gadE*, *hdeA* and *hdeB*, respectively, by qRT-PCR (Fig. 4.2). The upregulation of *gadX*, however, was not statistically significant in the microarray analysis, though the direction was the same. This discrepancy is possibly due to high inter-strain variation in expression within the genotype.

Expression of *dps*, which is involved in protecting DNA during starvation and acid stress (28), was upregulated by 3.6-fold in the bovine-biased genotype strains. Similarly, *clpA*, a chaperone necessary for protein degradation by the ClpAP protease (73), showed a 2-fold increase in expression. Other stress fitness-associated genes with increased expression in the bovine-biased genotype included the cold shock protein, *cspD* (173), cation transport regulators, *chaBC* (107) and the universal stress protein, *uspA* (26, 111) (Table 4.2). Moreover, expression of *katG* (108), *osmC* and *osmY* (168), the genes involved in resistance to peroxide and osmotic stress, also were upregulated

(Table 4.2). Interestingly, the general stress sigma factor, *rpoS* (169), was not differentially expressed between the two genotypes.

The RpoN regulon. Several metabolic genes, including nine genes involved in the nitrogen regulatory response that are regulated by the sigma factor RpoN (122), were upregulated in the bovine-biased genotype. The nitrogen regulatory protein, *glnK*, and the ammonium transporter, *amtB*, were both upregulated by 11.6- and 14.9-fold respectively. *glnD*, which is involved in the post transcriptional modification of *glnK*, also was upregulated in bovine-biased strains as were the nitrogen regulator I (*glnG*), a permease of the periplasmic glutamine binding protein (*glnH*), and genes associated with glutamate biosynthesis (*gltB*, *gltK* and *gltJ*) (Table 4.2).

Furthermore, *nac*, which encodes the nitrogen assimilation control protein, was upregulated by 12.7-fold in the bovine-biased genotype. Consequently, a number of Nac-regulated genes including *oppA*, *oppB*, *oppD*, *oppF*, *gabD*, *gabT*, and *dppA*, had higher expression levels in bovine-biased strains. Other RpoN-regulated genes such as *hypA*, *hypB* and *fdhF*, also had increased expression in the bovine-biased genotype (Table 4.2).

SNP genotyping and re-analysis of microarray data. Because the eight strains in this study were only characterized by the distribution of Stx insertion sites and represented the same multilocus sequence type (14), a more sensitive SNP genotyping method (88) was used to better understand the phylogenetic relationships of strains within and between the two genotypes. Among the four strains representing clinical genotype 1, three grouped together with a clade 8

control strain and one with a clade 1 control strain (Fig. 4.3). By contrast, all four strains representing the bovine-biased genotype 5 belonged to clade 7 (Fig. 4.3).

Since one clinical genotype 1 strain was part of a phylogenetically distinct lineage (clade 1) relative to the other three clinical genotype 1 strains (clade 8), the microarray data was re-analyzed after excluding the data generated from the clade 1 strain. The Fs test identified significant upregulation of 400 genes in the clinical genotype and 349 genes in the bovine-biased genotype (FDR <0.1, fold change >1.5) in this re-analysis. All but three genes (*bioB*, *dps* and *ycaL*) identified to be differentially expressed in the first analysis were also differentially expressed in the second analysis. Further, 561 additional genes were identified in the second analysis, as elimination of the clade 1 strain likely reduced the within-genotype variation. Twenty eight LEE genes including the genes encoding intimin (*eae*), the translocated intimin receptor (*tir*), and a positive regulator of LEE (*grlA*), were significantly upregulated in the three clinical genotype strains. Similarly, pO157-encoded genes such as *ehxA*, *toxB*, and 11 genes of the *etp* polycistron that encodes a T2SS, also were upregulated in the clinical genotype. As expected, the bovine-biased genotype strains showed increased expression of GAD and AFI genes relative to the three clinical genotype strains. One additional gene was *gadX*, which was significantly upregulated in the second analysis, but not the first analysis. The stationary phase sigma factor, *rpoS*, and *adiY*, which encodes an ARAC-like regulator of the arginine decarboxylase acid resistance system, also were upregulated in the bovine-biased genotype as were a number of stress fitness-associated and RpoN-regulated genes.

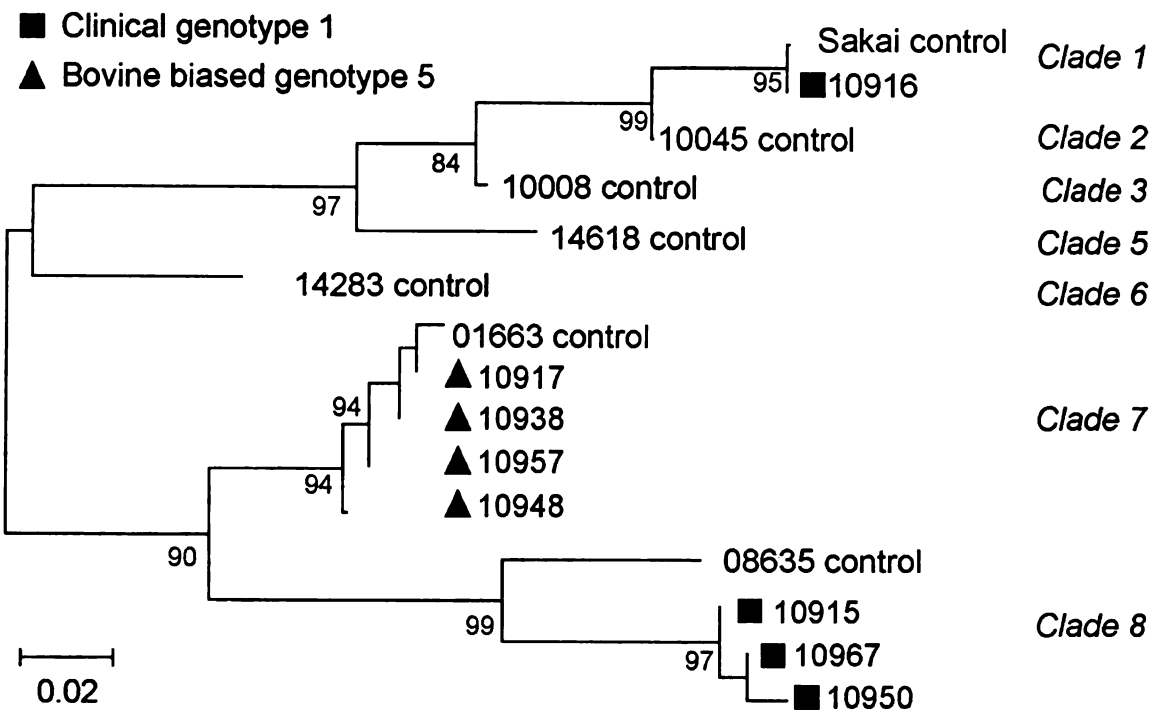


Figure 4.3. Neighbor joining phylogeny of single nucleotide polymorphism (SNP) genotypes representing the eight strains examined in the study. Three of the four clinical genotype 1 strains (black squares) belong to clade 8, whereas one clinical genotype 1 strain is part of clade 1. All four bovine-biased genotype 5 strains (black triangles) are members of clade 7.

Model stomach assay. To determine whether the increased expression of acid resistance genes in the bovine-biased genotype translates to a phenotypic difference, model stomach assays were conducted. These assays were used to directly compare the survival of both genotypes in a complex acidic environment that simulates the human stomach. Consistent with the microarray expression data, there was a significant difference ($P = 0.003$) between the survival rates of clinical and bovine-biased genotypes, as bovine-biased strains had a 2-fold increase in survival in the MSS (Fig. 4.4). The average survival rate per 30 min for the clinical genotype was -0.55±0.04, whereas the bovine-biased genotype was -0.27±0.04.

Figure 4.4. Survival of clinical and bovine-biased genotypes in the model stomach system. The average survival rate (log decrease in CFU/ml per 30 min) with the standard error of mean (SEM) from two independent experiments is plotted for each genotype.

DISCUSSION

Genotyping based on Stx-encoding bacteriophage insertion sites has demonstrated that the *E. coli* O157:H7 strains present in the bovine reservoir are considerably more diverse when compared to strains that cause human infections (14). Furthermore, it was suggested that some bovine-biased genotypes have reduced virulence and hence, cause disease less frequently relative to those bovine-derived clinical genotypes that are commonly isolated from patients (14). This variation could be due to gene content differences, including allelic variation in key genes among genotypes or due to expression differences in critical genes. Here, we describe differences in the expression of important genes that provide an explanation for the variation in infectivity between bovine-biased and clinical genotypes. Specifically, genome-wide expression profiling revealed differential expression of key virulence and stress fitness genes between the two genotypes, which was confirmed by qRT-PCR and a phenotypic assay. Because microarrays and qRT-PCR target different regions per gene, we suspect that the differential expression of genes identified in this study is not due to differences in gene content or allelic variation between genotypes, as both microarrays and qRT-PCR detected similar levels of expression.

One of the most important differences identified was the upregulation of the LEE in clinical versus bovine-biased genotypes. The LEE is considered a critical factor in *E. coli* O157:H7 disease pathogenesis, as it encodes a T3SS that mediates adherence to the intestinal mucosa (94). Because strains of the clinical

genotype expressed key LEE genes at a higher level, it is likely that these strains have an enhanced ability to adhere to the intestinal epithelium and cause the attaching and effacing lesions that initiate the disease process. By contrast, it is possible that increased expression of negative LEE regulators suppresses the expression of important LEE genes in the bovine-biased genotype, thereby reducing adherence and subsequent disease. The increase in *gadE* expression in the bovine-biased genotype supports this hypothesis, as our prior study determined that GadE, the central activator of the GAD system, negatively regulates LEE in O157:H7 strains (65). Similarly, GadX, a negative regulator of LEE in enteropathogenic *E. coli* (EPEC) (138), also was upregulated in the bovine-biased genotype as was *yhiF*, an AFI-encoded regulator that suppresses LEE expression in EHEC (156).

Similar to the LEE, another important factor in *E. coli* O157:H7 disease pathogenesis is the possession of the pO157 plasmid. The 92 kb plasmid carries genes for many putative virulence factors including *ehxA* (enterohemolysin) (129), *toxB* (toxin B) (19), and a number of *etp* genes necessary for a T2SS that secretes factors such as the StcE protease (44); all were upregulated in the clinical genotype. *ehxA* is a cell-associated, pore forming toxin from the repeats-in-toxin (RTX) family (129). Although the exact role of the enterohemolysin in O157:H7 pathogenesis is not clear, it has been shown to induce production of the interleukin-1beta (IL-1β) proinflammatory cytokine, which is a serum marker of HUS (152), and to cause injury to microvascular endothelium (5). The increased expression of *grlA*, a LEE-encoded positive regulator of *ehxA* (129),

may partly explain why *ehxA* was upregulated in the clinical genotype. StcE, a protease that is involved in the intimate adherence of EHEC to host cells, is secreted by the T2SS encoded by the *etp* gene cluster (44); eight of the *etp* genes were significantly upregulated in the clinical genotype. Similarly, *toxB*, which was upregulated in clinical genotype strains, also has been shown to be important for complete adherence to human epithelial cells (154). Together, these data demonstrate that the clinical genotype 1 strains are expressing factors important for adherence, suggesting that genotype 1strains may have an enhanced ability to adhere to human host cells and thus, are inherently more virulent than bovine-biased genotype 5 strains. A comparison of Stx expression was not possible as the genotypes used in this study do not have the same *stx* profiles.

During passage through the bovine gastrointestinal tract, *E. coli* O157:H7 has to survive a number of adverse environmental conditions, including extreme acidity in the abomasum (165), organic acid stress from volatile fatty acids in the rumen and colon, and occasional hyperosmolarity. Therefore, the capacity to withstand acidity and other environmental stresses is critical for O157:H7 strains to successfully persist in cattle. Consistent with this, we observed upregulation of acid resistance and stress fitness-associated genes in the bovine-biased genotype relative to the clinical genotype. The GAD system is the most efficient acid resistance system in *E. coli* (23, 75) and is essential for the survival of *E. coli* O157:H7 in bovines (119). All three GAD system genes (*gadA*, *gadB* and *gadC* (39)) had increased expression in the bovine-biased genotype. Moreover, AFI

genes, which are also involved in acid resistance (91), showed increased expression in the bovine-biased genotype. This increased expression of GAD and AFI genes is possibly induced by the upregulation of GadE, an essential activator of GAD and many AFI genes in *E. coli* O157:H7, as shown recently (65). DMEM buffered with MOPS maintains a neutral pH at exponential phase and therefore, it appears that under non-inducing conditions, the expression level of *gadE* and the GAD system genes is markedly higher in the bovine-biased genotype, which may enhance survivability in the bovine gastric environment (pH 2.1) (165). This inference was confirmed by the model stomach assays, where bovine-biased strains survived better than clinical genotype strains. The model stomach represents a complex acidic environment and hence, the increased survivability of bovine-biased strains may also be attributable to the increased expression of stress fitness genes such as *dps*, *clpA* and *uspA*, in addition to the acid resistance genes. Together, these findings indicate that strains of the bovine-biased genotype 5 are more resistant to adverse environmental conditions, which likely facilitates survival and the subsequent colonization of bovines. A prior study observed similar results by comparing resistance to acetic acid (pH 3.3) among *E. coli* O157:H7 strains isolated from environmental sources and humans (103). Specifically, those isolates from humans were found to be less resistant to acetic acid than isolates obtained from bovine feces (103). Our study suggests that this difference in acetic acid resistance could be due to differences in expression of important acid and stress resistance genes. The variation in resistance to acidity among different genotypes within the bovine

population is interesting, and we demonstrate that those bovine-derived strains that can infect humans are actually less resistant to acid than those strains that colonize the bovine. This characteristic enables less resistant strains, or genotypes, to express alternative factors, particularly those involved in adherence, when in contact with human cells mainly due to the negative interaction between acid resistance regulators and adherence genes. Although bovine-biased genotypes that are highly resistant to acid can get transmitted to humans, these strains cause disease much less frequently, which could possibly be due to the decreased expression of important virulence factors.

As the nitrogen source in DMEM is glutamine, strains face an ammonia limiting condition during growth in DMEM. This condition leads to the induction of the nitrogen regulatory response (122), and expression data from this study indicates that bovine-biased strains are more efficient at mounting this response at the transcriptional level. Ten RpoN-regulated genes involved in the response were upregulated in bovine-biased strains, and important regulators and assimilation proteins such as *glnK*, *amtB* and *nac*, had more than a 10-fold increase in expression. These observations indicate that strains of the bovine-biased genotype are equipped with a more efficient nitrogen regulatory response system, which can enhance survival in ammonia limiting conditions. Such conditions are likely encountered in the bovine gastrointestinal tract or external environment.

A recent study from our laboratory classified clinical O157:H7 strains into nine clades based on a highly sensitive SNP genotyping method (88). The

severity of disease was shown to vary between the clades, with clade 8 being associated with HUS. Interestingly, three of the four clinical genotype 1 strains used in this study belonged to clade 8, thereby demonstrating the presence of this clade in the bovine reservoir. After excluding the phylogenetically distinct clinical genotype 1 strain, the three clade 8 strains still had increased expression of several LEE and pO157 genes. In contrast, all of the bovine-biased strains used in this study belonged to clade 7, a lineage that was associated with less severe clinical disease (88).

Decreasing the prevalence of O157 colonization in cattle has become increasingly significant in the current strategies to control O157 infections, as it is less feasible to completely prevent fecal contamination of food vehicles such as vegetables, fruit juices and beef (34). In this context, identifying the genes that are critical for O157 persistence in cattle is important for developing novel prevention strategies against this pathogen. This study, along with previous studies (119), indicates that acid and stress resistance genes encompass an important set of genes that are crucial for survival of *E. coli* O157:H7 in cattle and hence, can be used as targets for prevention measures. In addition, there is considerable variation in the expression of different genes between genotypes of *E. coli* O157:H7 isolated from cattle, with strains of clinical genotype 1 having increased expression of important virulence factors. The upregulation of key virulence components provides a possible explanation for the predominance of this genotype in clinical cases.

ACKNOWLEDGEMENTS

This project has been supported by Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No. N01-AI-30055.

The authors wish to thank Lindsey M. Ouellette and Hans Steinsland for technical assistance, Teresa M. Bergholz for helpful suggestions for the microarray design, and Linda S. Mansfield, Galeb S. Abu-Ali and James T. Riordan for critical review of the manuscript.

CHAPTER 5

Summary and future directions

SUMMARY

Acid resistance (AR) is a critical factor in the pathogenicity of *E. coli* O157:H7 as it confers the ability to be transmitted through a variety of food vehicles and to breach one of the first host defense barriers—the gastric acidity—while passing through the human stomach. *E. coli* O157:H7 has been shown to be more efficient in resisting complex acidic conditions when compared to other strains of *E. coli* (10). This superior AR is considered as the main contributing factor for the low infective dose of *E. coli* O157:H7. Four distinct mechanisms are known to contribute to the AR of *E. coli* including the OXI system, GAD system, ARG system, and LYS system, of which GAD system is most efficient (39, 75). The GAD system is comprised by two glutamate decarboxylase isozymes, GadA and GadB, and a GABA-glutamate antiporter, GadC. The GAD system is one of the most complexly regulated systems in *E. coli*, which indicates the importance of this system in the maintenance of normal physiology of the organism (39). Furthermore, animal experiments have shown that a functional GAD system is necessary for the survival of *E. coli* O157:H7 in cattle (119).

Maximum GAD expression is induced at stationary phase of growth or when the cells are exposed to acidic pH (12, 39). At least 14 regulators are involved in the regulation of GAD system, among which GadE functions as the central activator of the system. A functional GadE is essential for the expression of GAD components at all growth phases in rich or minimal media (54, 84). GadE is a DNA binding LuxR family regulator and it induces GAD expression by directly

binding to the 20 bp GAD box region upstream of *gadA* and *gadBC* genes. Even though the GAD system and its regulation are well characterized in benign laboratory strains of *E. coli*, these aspects remain largely unknown in pathogenic strains of *E. coli*. Most importantly, the function of GadE at a genome wide scale is not known for any of the pathogenic strains of *E. coli*. *E. coli* O157:H7 genome is markedly different from the laboratory strains of *E. coli* as it has acquired many foreign DNA elements through horizontal gene transfer during evolution. It is possible that during the process of integrating laterally acquired DNA elements into the chromosomal regulatory network, a chromosomal regulator such as GadE has evolved into a global regulator with multiple functions. Therefore, I hypothesized that GadE has additional functions besides the positive regulation of GAD system and the purpose of the first part of this project was to identify the GadE regulon and to characterize the AR phenotype of a *gadE* mutant *E. coli* O157:H7.

The *E. coli* O157:H7 Sakai strain isolated from the radish sprout outbreak in Japan was used in this part of study (46, 96). Here, the experimental approach for identifying the GadE regulon was to compare the transcription profiles of wild type (WT) and Δ*gadE* strains at different growth phases to detect the genes whose expression is affected by the inactivation of *gadE*. An *E. coli* O157:H7 Δ*gadE* strain was constructed by a one step inactivation method (32, 98). Inactivation of *gadE* did not affect the growth patterns of *E. coli* O157:H7. WT and mutant strains were grown in MOPS minimal medium up to exponential and stationary phase and the transcriptomes were compared using whole genome

microarrays. Inactivation of *gadE* affected the expression of 60 genes independent of growth phase and 122 genes in a growth phase-dependent manner. Expression of *gadE* has been shown to be growth phase-dependent (12) and hence, this growth phase effect on genes was expected. Additionally, putative GadE binding sequences or GAD boxes were observed upstream of eight genes that had significant differential expression due to *gadE* inactivation indicating a direct regulation of these genes by GadE.

Genes involved in AR such as the GAD and AFI genes comprised an important part of the GadE regulon. Expression of all the three components of the GAD system, *gadA*, *gadB* and *gadC,* was markedly decreased by the inactivation of *gadE*. Similarly, *hdeBAD*, the AFI genes involved in resistance to self metabolites and AR at high cell density showed decrease in expression in Δ*gadE*. These findings were similar to that in the laboratory strain, *E. coli* K-12 in which inactivation of *gadE* completely abolished the expression of GAD and markedly reduced the expression of AFI genes (84, 91). Here, we also found that the effect of *gadE* inactivation on GAD and AFI genes was significantly higher at stationary phase than at exponential phase. The regulation of GAD and AFI genes by GadE appears to be direct because GAD boxes with 100% similarity were detected upstream of *gadA* and *gadBC* and two putative Gad boxes were identified upstream of *hdeD*.

The acid resistance phenotype of Δ*gadE* was characterized using functional assays such as AR mechanism assays and a model stomach assay. AR mechanism assays for GAD, ARG and OXI systems assess the effect of

gadE inactivation on the functioning of the three systems. As expected, there was complete abrogation of GAD system functioning in the Δ*gadE*. Interestingly, inactivation of *gadE* also affected the ability of ARG and OXI systems to protect the strains from acidic challenge. Complementation with wild type *gadE* restored the phenotype for GAD and ARG systems, but not for the OXI system. We believe that flooding the cells with wild type *gadE* through a multi copy plasmid might have negatively affected the OXI system. In the model stomach system, inactivation of *gadE* abolished the ability of *E. coli* O157:H7 to survive in this complex acidic environment. The Δ*gadE* did not survive for even 30 min in the model stomach whereas wild type and complement survived up to 90 min with minimal decrease in cell counts. This demonstrates that *gadE* is essential for the survival of *E. coli* O157:H7 in a simulated gastric environment.

A previous study in *E. coli* O157:H7 Sakai strain has shown an increased adherence phenotype for Δ*gadE*. Expression analysis of WT and Δ*gadE* from the same study found that inactivation of *gadE* resulted in increased expression of LEE4 genes (155). However, the mechanism underlying this negative regulation was not identified as the expression of Ler, the key positive regulator of LEE, remained unchanged in the Δ*gadE* (95). Our study found that GadE negatively regulates genes of all five of the LEE operons including *ler*. To determine the mechanism by which GadE negatively regulates LEE, a Δ*gadE*Δ*ler* strain was constructed and the expression of LEE genes was analyzed. Inactivation of *ler* reversed the effect of *gadE* inactivation on LEE indicating that GadE acts through

Ler. Presence of a putative GAD box upstream of *ler* also suggested that GadE directly suppresses *ler* expression resulting in decreased LEE expression.

Expression of LEE is affected by several environmental factors such as temperature, osmolarity and bicarbonate ion concentration in the medium (2, 157, 164). However, the effect of acidity on the LEE is not well investigated. It is important to understand how acidic pH affects a major virulence factor of *E. coli* O157:H7 such as LEE as it may provide insights into the temporal expression of virulence factors as the organism passes through the stomach. As GadE is a negative regulator of LEE and as acidic pH induces GadE, LEE expression possibly decreases at low pH. Supporting this, there was a 6 – 9-fold decrease in expression of LEE genes including, *ler*, when *E. coli* O157:H7 cells were grown at pH 5.0, a moderate acidic condition. GadE appears to control this decrease in expression of *ler* as there was no reduction in *ler* expression in the Δ*gadE* at pH 5.0. However, the other LEE genes tested still showed a partial decrease in expression in the Δ*gadE* at acidic pH indicating that other regulators may also be involved in this mechanism.

Other GAD regulators such as GadX and EvgA have been shown to negatively regulate LEE expression in EPEC (99, 139). GadX negatively regulates LEE expression through the PerA regulator at pH 5.5 in a complex medium (139). To determine whether the same mechanism exists in *E. coli* O157:H7, a Δ*gadX* strain was constructed and the LEE expression was analyzed at different growth conditions. Unlike in EPEC, inactivation of *gadX* did not have a marked positive effect on the expression of LEE genes at acidic conditions in

E. coli O157:H7. The PerA regulator is not present in *E. coli* O157:H7 (61) and that may be the reason for the lack of GadX regulation of LEE. On the other hand, *evgA* inactivation in *E. coli* O157:H7 resulted in a marked upregulation of the LEE genes *tir*, *espD*, *eae*, and *sepZ* demonstrating that similar to EPEC (99), a negative interaction between EvgA and LEE exists in EHEC as well.

Subsequently, to determine the effect of gastric acidity on the expression of the major virulence factors of *E. coli* O157:H7, stationary phase cells were exposed to pH 2.0 and the expression of virulence factors were analyzed. Exposure to pH 2.0 resulted in a marked decrease in the expression of virulence genes, *stx1*, *stx2*, LEE genes and pO157 genes. On the other hand, there was an increase in expression of the three GAD system genes, *gadA*, *gadB* and *gadC*. These results indicate that upon exposure to extreme acidity, *E. coli* O157:H7 upregulates the expression of AR genes that are required for protection from acidity and concurrently downregulates the expression of virulence genes whose expression is not required in such an environment. This appears to be an efficient survival strategy because *E. coli* O157:H7 is not a gastric pathogen and therefore, expressing the virulence factors in stomach is unnecessary and may cause a heavy burden on the energy conservation and survival of the organism.

Identification of differentially expressed genes between *E. coli* O157:H7 genotypes

E. coli O157:H7 strains isolated from clinical cases and from cattle can be broadly divided into two genotypes, clinical and bovine-biased. Clinical genotypes predominate in clinical cases, but are also isolated from bovine sources. On the other hand, bovine-biased genotypes are isolated mostly from bovine sources and rarely from clinical cases (14). We hypothesized that this variation in infectivity of the two genotypes is due to differential expression of virulence and stress fitness-associated genes.

To test this hypothesis, four strains were selected from each genotype and their exponential phase expression profiles were compared. A large number of genes were differentially expressed between the two genotypes. Importantly, major virulence factors of *E. coli* O157:H7 such as LEE and pO157 plasmid encoded genes had increased expression in clinical genotype strains. On the other hand, genes that are essential for acid and stress resistance had higher expression in bovine-biased genotype strains. All of the three components of the GAD system, *gadA*, *gadB* and *gadC,* and many of the AFI genes were upregulated in strains of the bovine-biased genotype. The central activator of GAD system, *gadE*, also had higher expression in bovine-biased genotype, which could be the reason for increased expression of GAD and AFI genes. The increased expression of *gadE* could have contributed to the decreased expression of LEE in the bovine-biased genotype because as we demonstrated in these studies GadE is a negative regulator of LEE in *E. coli* O157:H7 (65,

155). This increased expression of AR genes in strains of the bovine-biased genotype was confirmed functionally by a model stomach assay in which bovine-biased strains demonstrated at least twice as much survival rate as the clinical genotype strains. Also upregulated in strains of the bovine-biased genotype were the genes of the RpoN regulon. These genes are involved in the nitrogen regulatory response in ammonia limiting conditions (122) and hence, their increased expression could contribute to better survival in a nutrient limiting environment.

Single nucleotide polymorphism (SNP) genotyping of the clinical and bovine-biased genotype strains used in this study confirmed that they belong to distinct genetic populations. Three of the clinical genotype strains belonged to clade 8, which has been shown to be associated with more severe disease including a higher incidence of HUS (88) and one strain belonged to clade 1, which has also been frequently isolated from clinical cases (88, 96). Bovine-biased genotype strains on the other hand belonged to a clade that is less frequently isolated from clinical cases, clade 7 (88).

Together, the findings from this study suggest that one of the reasons for the variation in distribution of *E. coli* O157:H7 genotypes between clinical cases and cattle could be the differences in expression levels of genes involved in virulence and resistance to environmental conditions. Presence or absence of genes could also lead to variation in infectivity; however, we have not addressed that aspect in this study. Based on the results from this study, we hypothesize that increased expression of virulence factors such as LEE and pO157 genes is

the basis for the predominance of clinical genotype strains in human cases. In contrast, bovine-biased strains are more resistant to acid and other environmental stress, which helps them persist in the bovine reservoir. At the same time, due to the negative interactions between acid resistance and virulence genes, bovine-biased genotype strains express virulence factors at a lower level, which may be the reason for their decreased prevalence in clinical cases. It is possible that these highly environmental stress-resistant bovine-biased strains reach the humans through various food vehicles, but due to decreased expression of important adherence and virulence factors do not cause disease in humans. This study also demonstrated that clinical and bovine-biased genotype strains belong to distinct SNP genotypes or clades, which vary in their ability to cause disease in humans. Presence of clade 8 strains among the clinical genotype strains isolated from cattle supports the argument that clade 8 outbreak strains such as the spinach strain could have originated from a cattle farm. Moreover, it demonstrates that SNP genotyping could be used as a technique for detecting highly virulent strains in the reservoir host itself, which may help in prevention and control of *E. coli* O157:H7 outbreaks.

FUTURE DIRECTIONS

Effects of *gadE* inactivation on the survival and virulence of *E. coli* O157:H7 in vivo

An important finding from this study was that inactivation of *gadE* in *E. coli* O157:H7 resulted in increased expression of important adherence factors that are needed for colonization of the host. Furthermore, a functional *gadE* was found to be necessary for the survival of *E. coli* O157:H7 in a simulated gastric environment. Naturally, the next step is to assess the survival and virulence of Δ*gadE* in vivo using the best animal models of the human disease. Based on the expression results, we hypothesize that even though *E. coli* O157:H7 Δ*gadE* expresses virulence factors such as LEE at a higher level, it may not be able to cause disease because of its inability to cross the gastric acid barrier in the absence of GadE.

However, lack of a good animal model has been an issue hindering in vivo studies for *E. coli* O157:H7. Several animal models such as rabbits, conventional mice, specific pathogen free (SPF) mice and streptomycin-treated mice have been used as disease models. However, in most of them development of clinical signs and pathological lesions are only minimal (97, 124). Recently, germ free Swiss-Webster mice have been shown to develop HUS after EHEC infection even with infectious dose as low as 100 cells. In these germ free mice, the bacteria were found adherent to the cecal and ileal mucosa and caused renal lesions unlike other animal models (35). Therefore, germ free mice could be used

for assessing the pathogenicity of Δ*gadE E. coli* O157:H7. A caveat in using a mouse model for determining the survival of Δ*gadE* is that the gastric pH of mice is significantly higher than that of the human stomach (92). The mouse stomach pH varies between 3.0 (fed) and 4.0 (fasted) (92) and hence, the Δ*gadE* may be able to survive better and the results may not be representative of the conditions when bacteria must pass through the human stomach.

An alternative approach is to use cattle as animal models for assessing the survival of Δ*gadE*. Cattle are the reservoir hosts of *E. coli* O157:H7 and the pH of cattle abomasum (the true stomach) is similar to that of the human stomach (165). Cattle could provide an ideal environment to determine the role of GadE in survival in gastric acid in vivo. However, the pathogenicity of Δ*gadE* cannot be assessed in cattle as O157 does not cause any clinical disease in them. Also the presence of a four chambered stomach including the rumen in cattle may affect the results. Alternately, the rabbit animal model could be used, which has a stomach pH of 1 – 2 (37). However, similar to some of the mouse models, rabbit is not a well-established model for clinical symptoms for *E. coli* O157:H7. Therefore, assessing the survival and pathogenicity of Δ*gadE E. coli* O157:H7 in the same animal model appears to be a less feasible option currently.

GAD system in *E. coli* O157:H7

This study, along with some of the previous studies from our lab (8, 10) demonstrates that there are many differences in regulation and functioning of GAD system in *E. coli* O157:H7 compared to the laboratory strain *E. coli* K-12, in which most of the studies on GAD system have been conducted. However, many aspects of the GAD system in *E. coli* O157:H7 still remain unclear. Most importantly, the contribution of GadA and GadB to the AR of O157 has not been investigated. Specifically, it is not known whether one of these genes is sufficient or both are necessary for AR of O157. In *E. coli* K-12, at pH 2.0 both *gadA* and *gadB* are required for AR whereas at higher pH such as 2.5 either one of them can provide protection (23). This aspect could be different in O157 because unlike other *E. coli* strains *gadA* and *gadB* sequences remain divergent with distinct regulatory regions in O157 (8), which might allow them to function more independently. Also, *E. coli* O157:H7 has been shown to have superior AR compared to other *E. coli* strains (10). Hence, a study comparing the survival of Δ*gadA* and Δ*gadB E. coli* O157:H7 strains at pH 2.0 – 2.5 in minimal and complex environments is needed to understand the functioning of GAD system in O157. I hypothesize that under a minimal acidic environment where the only stress is acidity, presence of either *gadA* or *gadB* is sufficient for survival of *E. coli* O157:H7. However, in a complex acidic condition such as model stomach, which presents multiple stresses, both these genes may be essential for maximum survival.

Two conditions that induce the GAD system are stationary phase of growth and acidic pH. Stationary phase induction of GAD system is well investigated and it is known that the alternative sigma factor, RpoS regulates the GAD expression through GadXW at the stationary phase (39). However, the signaling pathways and regulators that induce the GAD system at acidic pH, especially at exponential phase, remain unknown. It has been shown that low pH conditions lower the physiological concentration of cAMP and CRP. This reduction in cAMP-CRP levels can initiate increased transcription of RpoS, which can then induce *gadX* and *gadE* transcription leading to increased expression of *gadABC* (86). However, the expression of RpoS is minimal at exponential phase (39) and therefore I hypothesize that other unidentified regulators are also involved in the acidic induction of GAD system at exponential phase. To address this, a transposon mutagenesis approach could be used in which the survival of mutants at acidic pH during exponential phase is assessed. This might help to identify the regulators that are involved in the induction of the GAD system at acidic pH.

Interactions between GadE and Ler in *E. coli* O157:H7

Comparison of *E. coli* O157:H7 wild type and Δ*gadE* expression profiles showed that GadE is a negative regulator of LEE and a functional Ler is essential for this negative regulation. Presence of a putative GAD box region upstream of *ler* implied that GadE directly binds to this GAD box and repress the *ler* expression. However, this binding was not confirmed experimentally. An electrophoretic mobility shift assay (EMSA) with tagged-GadE protein and a DNA fragment containing the putative GAD box upstream of *ler* could confirm whether the *ler* GAD box is a functional GadE binding region. Further, once confirmed as a GadE binding region, this GAD box could be inactivated in a *gadE*-over expressing strain to demonstrate that in the absence of the *ler* GAD box, GadE mediated repression of *ler* and LEE does not occur.

Differences between clinical and bovine-biased genotypes of *E. coli* O157:H7

Numerous virulence and acid/stress resistance genes were found to be differentially expressed between strains of the clinical and bovine-biased genotypes of *E. coli* O157:H7 in this study. Increased expression of AR genes in the bovine-biased genotype was confirmed functionally by model stomach assays. However, phenotypic assays have not been done to confirm that the statistically significant upregulation of LEE genes and pO157 encoded virulence genes in the clinical genotype is biologically significant.

I hypothesize that increase in expression of adherence factors such as LEE in clinical genotype can lead to an increase in adherence of these strains to the epithelial cells. This could be confirmed by epithelial cell association assays, which measure the adherence and invasion of bacteria to the epithelial cell. In our lab, bovine mammary epithelial cells (MAC-T) have been used extensively to assess the association of *E. coli* O157:H7 with host cells and the same could be used for comparing clinical and bovine-biased genotypes. Another approach would be to quantify the association of bacteria with epithelial cells using flow cytometry, in which the bacteria would be labeled with CFDA-SE dye and the labeled bacteria would be used to infect the MAC-T cells. The level of fluorescence from MAC-T cells, as assessed by flow cytometry, will indicate levels of adherence/invasion of *E. coli* O157:H7.

Enterohemolysin (Ehx) assay, which quantifies the hemolytic activity of *E. coli* O157:H7 in defibrinated sheep blood (129) could be used to confirm the

increased expression of pO157-encoded hemolysin (*ehxA*) in clinical genotype strains.

Finally, animal model experiments comparing the pathogenicity of clinical and bovine-biased genotypes are needed to confirm the results of expression studies and phenotypic assays. Germ free mice, which have been recently established as a model for *E. coli* O157:H7-induced disease, could be used for this purpose.

SNP genotyping of the strains used in this study showed that clinical and bovine-biased strains belong to distinct genetic populations of *E. coli* O157:H7. Clinical genotype strains belonged to hyper virulent clade 8 and clade 1 whereas bovine-biased strains were all clade 7, a group that causes human disease less frequently (88). An extension of these findings would be to conduct SNP genotyping of a larger number of bovine *E. coli* O157:H7 isolates that belong to clinical and bovine-biased genotypes and analyze the distribution of clades between them. If the clinical genotype strains isolated from bovines belong to groups that cause severe disease such as clade 8 and clade 2, that could provide further explanation to the predominance these genotypes in human cases. This will also validate SNP genotyping as a superior method for identifying virulent *E. coli* O157:H7 strains in clinical settings and cattle farms.

APPENDICES

APPENDIX 1

R/MAANOVA: steps in expression data analysis

1. Preparing the data file

Assemble the data from all the arrays into one MS excel file that contains the following information; i) location information for a spot including grid, metarow, metacolumn, row and column ii) probe ID iii) data from each array in the order, Cy3, Cy5 and flag. Save the file in TAB delimited text format.

2. Preparing the design file

This file contains information about the microarray experimental design. It is prepared in MS excel. The number of rows depends on the number of arrays and number of columns depends on the number of factors that are included in the design. The file should include information about array, dye and sample. Sample indicates the biological replicate. Also included in design file is information about the factors included in the experimental design such as strain, growth phase, and treatment. Save the file in TAB delimited text format.

3. Launch R and load MAANOVA

4. Load the data

Use read.madata function.

5. Normalize the data

Use transform.madata function to normalize the data by spatial-intensity joint lowess (rlowess).

6. Create the data

The final data set used for statistical analysis is created by createData function by collapsing the replicates and log transforming the values.

7. Make the model

Make a statistical model for the data analysis using makeModel function. Include all the factors in the formula such as array, dye, sample, strain, and growth phase in the same order as in the design file and mention which are the random factors.

8. Fit the model into the data

Use fitmaanova function. Here the statistical model is fitted into the data for each gene.

9. Test for differential expression

Use matest function. Significant differences in expression are identified by F test. Four types of F tests are available in maanova among which Fs test is considered to provide the best results as it does not assume variance is homogenous across genes. P-values are calculated by permutation test.

10. Correction for multiple comparisons

Use adjPval function. This step calculates FDR adjusted p-values for the test results by Benjamini-Hochberg linear step-up correction.

APPENDIX 2

TABLE 2.S1: Genes showing significant strain effect (FDR < 0.1)

ECs number[a]	Gene[b]	Function	Expression ratio[c] (Δ*gadE*/ WT)	Expression ratio[d] (Stat/exp)
Genes significantly down-regulated in Δ*gadE*				
ECs 2294		putative oxidoreductase, major subunit	0.75	1.8
ECs 3904		putative transport periplasmic protein	0.82	0.5
Genes significantly up-regulated in Δ*gadE*				
ECs 0323	*yagZ*	hypothetical protein	1.32	2.0
ECs 1114	*nusA*	partial putative tail component of prophage CP-933R	1.22	1.8
ECs 1305	*yjcT*	putative helicase	1.19	3.2
ECs 1471	*fabG*	3-oxoacyl-[acyl-carrier-protein] reductase	1.25	1.14
ECs 1830	*yciF*	putative structural proteins	1.31	6.2
ECs 1994	*nleG2-2*	non-LEE effector protein	1.26	2.5
ECs 2060	*vgrE* *	unknown protein associated with Rhs element, VgrE protein	1.24	2.2
ECs 2156	*nleG2-3*	non-LEE effector protein	1.28	2.9
ECs 4262		hypothetical protein	1.26	0.5
ECs 4270	*cysG*	hypothetical membrane protein	1.29	0.6
ECs 4271	*recG*	putative ATP binding protein of ABC transporter	1.29	1.3
ECs 4272	*pdxB*	hypothetical membrane protein	1.32	3.5
ECs 4280	*ilvG*	putative DNA processing chain A	1.30	3.3
ECs 4300	*hcaT*	putative membrane protein	1.43	7.9
ECs 4307	*basS*	Signal transduction mechanisms	1.43	5.4
ECs 4326	*yeeA*	unknown function	1.22	0.2
ECs 4327		putative phospholipid biosynthesis acyltransferase	1.17	0.3
ECs 4334	*rhaB*	hypothetical protein	1.23	0.6
ECs 4335	*yegB*	putative membrane/ transport protein	1.43	0.6

ECs number[a]	Gene[b]	Function	Expression ratio[c] (Δ*gadE*/ WT)	Expression ratio[d] (Stat/exp)
ECs 4336	*exuT*	hypothetical membrane protein	1.32	0.6
ECs 4337	*ynhE*	hypothetical lipoprotein	1.28	0.6
ECs 4338	*wrbA*	putative 3-oxoacyl-synthase II	1.27	0.8
ECs 4339		putative beta-hydroxydecanoyl-ACP dehydrase	1.22	0.8
ECs 4382	*chuT*	putative hemin binding protein	1.39	7.6
ECs 4393	*yhiU*	member of acid fitness island, component of the MdtEF multidrug transporter	1.46	1.5
ECs 4550	*espF*	type III secretion system, secreted effector protein	1.30	0.6
ECs 4553	*cesD2*	type III secretion system chaperone	1.36	0.5
ECs 4555	*espD*	type III secretion system, secreted translocator protein	1.38	
ECs 4558	*escD*	type III secretion system, structural protein	1.18	1.3
ECs 4560	*cesT*	type III secretion system, chaperone	1.38	0.3
ECs 4561	*tir*	translocated intimin receptor protein	1.39	0.5
ECs 4562	*map*	type III secretion system, secreted effector protein	1.26	
ECs 4563	*cesF*	type III secretion system, chaperone	1.31	1.6
ECs 4564	*espH*	type III secretion system, secreted effector protein	1.33	3.1
ECs 4565	*sepQ*	type III secretion system, structural protein	1.27	2.9
ECs 4567	*orf15*	orf of unknown function	1.19	0.9
ECs 4571	*sepZ* *	type III secretion system, secreted effector protein	1.36	6.4
ECs 4572	*rorf8*	orf of unknown function	1.26	3.0
ECs 4575	*escC* *	type III secretion system, structural protein	1.21	0.3
ECs 4584	*orf5*	orf of unknown function	1.34	0.5
ECs 4585	*orf4*	orf of unknown function	1.41	0.6
ECs 4586	*orf3*	orf of unknown function	1.37	0.6
ECs 4587	*cesAB*	type III secretion system, chaperone	1.35	0.6
ECs 4588	*ler* *	type III secretion system, regulator	1.35	0.8

ECs number[a]	Gene[b]	Function	Expression ratio[c] (Δ*gadE*/ WT)	Expression ratio[d] (Stat/exp)
ECs 4590	*espG*	type III secretion system, secreted effector protein	1.27	0.6
ECs 5072		putative carbohydrate ABC transport system permease	1.23	0.5
ECs 5074	*lysA*	putative histidine protein kinase	1.30	0.5
ECs 5110	*yjdK*	lysine tRNA synthetase, inducible; heat shock protein	1.39	0.8
ECs 5218	*treR*	repressor of treA,B,C	1.28	4.8
ECs 5242		putative integrase	1.31	0.3
ECs 5248		orf of unknown function	1.28	1.2
ECs 5249		putative resolvase	1.23	1.2
ECs 5256		orf of unknown function	1.28	0.5
ECs 5257		orf of unknown function	1.23	0.6
ECs 5305		orf of unknown function	1.21	
ECs 5532		orf of unknown function	1.26	0.7
pO157p38		plasmid gene	1.23	1.4
pO157p80		plasmid gene	1.23	0.09

[a] Locus ID for *E. coli* O157:H7 Sakai strain (Genbank # BA000007).

[b] The genes marked with asterisks have putative GAD boxes upstream of their sequence.

[c] Expression ratio = $2^{[\log2(\Delta gadE) - \log2(WT)]}$

[d] Expression ratio between two growth phases = $2^{[\log2(stat) - \log2(exp)]}$ determined by microarray, ratios are reported only for genes with a significant growth phase effect (FDR<0.05)

APPENDIX 3

TABLE 2.S2: Genes with significant interaction between growth phase (exponential and stationary) and strain (wild type and Δ*gadE*) effects (FDR< 0.05)

ECs number	Gene[a]	Function	WT/ Δ*gadE* Exp phase	WT/ Δ*gadE* Stat phase
ECs0017	*nhaA*	putative cell division protein	1.2	0.7
ECs0120	*lpdA*	lipoamide dehydrogenase (NADH); component of 2-oxodehydrogenase and pyruvate complexes	1.1	0.8
ECs0282		unknown protein from prophage CP-933H	1.1	0.8
ECs0283	*lyxK*	unknown protein from prophage CP-933H	1.1	0.9
ECs0345	*ykgF*	hypothetical protein	1.2	0.8
ECs0464	*tsx*	nucleoside channel; receptor of phage T6 and colicin K	1.1	0.8
ECs0483	*cyoD*	cytochrome o ubiquinol oxidase subunit IV	1.2	0.8
ECs0484	*cyoC*	cytochrome o ubiquinol oxidase subunit III	1.2	0.8
ECs0681	*ybeL*	putative alpha helical protein	0.8	1.3
ECs0702	*yleB*	putative RNA	0.8	1.4
ECs0746	*sdhC*	succinate dehydrogenase, cytochrome b556	1.2	0.8
ECs0747	*sdhD*	succinate dehydrogenase, hydrophobic subunit	1.2	0.8
ECs0748	*sdhA*	succinate dehydrogenase, flavoprotein subunit	1.2	0.8
ECs0749	*sdhB*	succinate dehydrogenase, iron sulfur protein	1.2	0.7
ECs0750	-	hypothetical protein	1.1	0.7
ECs0751	*sucA*	2-oxoglutarate dehydrogenase (decarboxylase component)	1.2	0.7

ECs number	Gene[a]	Function	WT/ ΔgadE Exp phase	WT/ ΔgadE Stat phase
ECs0752	*sucB*	2-oxoglutarate dehydrogenase (dihydrolipoyltranssuccinase E2 component)	1.1	0.7
ECs0759	*htrE*	unknown function	1.1	0.8
ECs0804	-	unknown protein encoded by prophage CP-933K	1.1	0.9
ECs0881	*ybiI*	hypothetical protein	0.7	1.5
ECs1012	*ompF*	outer membrane protein 1a (Ia;b;F)	1.1	0.8
ECs1028	*ycbF*	putative chaperone	1.1	0.9
ECs1122	-	partial putative outer membrane protein Lom precursor encoded by prophage CP-933R	1.1	0.9
ECs1257	-	putative synthetase	0.9	1.3
ECs1271	*ycdT*	hypothetical protein	1.2	0.8
ECs1355	*terD*	putative tellurium resistance protein TerD	1.1	0.8
ECs1356	*terE*	putative phage inhibition, colicin resistance and tellurite resistance protein	1.1	0.8
ECs1377	*yjiM*	unknown function	1.1	0.9
ECs1490	*ycfR*	hypothetical protein	0.8	1.3
ECs1663	*ompT*	outer membrane protein 3b (a), protease VII	1.2	0.7
ECs1668	*minE*	cell division topological specificity factor, reverses MinC inhibition of ftsZ ring formation	1.1	0.8
ECs1669	*minD*	cell division inhibitor, a membrane ATPase, activates minC	1.1	0.8
ECs1711	*ychM*	hypothetical protein	0.7	2.1
ECs1722	*chaB*	cation transport regulator	0.6	3.2
ECs1744	*oppB*	oligopeptide transport permease protein	1.2	0.8
ECs1825	*yicL*	unknown function	1.1	0.8

ECs number	Gene[a]	Function	WT/ ΔgadE Exp phase	WT/ ΔgadE Stat phase
ECs1829	*yciE*	hypothetical protein	1.2	0.7
ECs1879	*goaG*	4-aminobutyrate aminotransferase	0.8	1.4
ECs1880	*pspF**	psp operon transcriptional activator	0.6	2.9
ECs1902	*tyrR*	putative acyltransferase	0.5	3.4
ECs2018	*ygiB*	unknown function	1.1	0.9
ECs2061	-	unknown protein associated with Rhs element	1.1	0.8
ECs2097	*gadC**	Glutamate-GABA antiporter	0.5	4.1
ECs2098	*gadB**	glutamate decarboxylase isozyme	0.4	8.0
ECs2099	*pqqL*	putative peptidase	0.8	1.5
ECs2100	*yddB*	hypothetical protein	0.7	2.0
ECs2260		unknown protein encoded by cryptic prophage CP-933P	0.5	4.3
ECs2279	*mdoB*	putative repressor protein encoded by cryptic prophage CP-933P	0.8	1.2
ECs2432	-	hypothetical protein	0.9	1.2
ECs2453	-	hypothetical protein	0.8	1.4
ECs2514	*fadD*	acyl-CoA synthetase, long-chain-fatty-acid--CoA ligase	0.8	1.4
ECs2613	*ftn*	cytoplasmic ferritin (an iron storage protein)	1.2	0.8
ECs2692	-	hypothetical protein	0.8	1.3
ECs2839	-	GDP-mannose dehydratase	1.2	0.8
ECs2840	*cchA*	glycosyl transferase	1.1	0.8
ECs2841	*ydeB*	perosamine synthetase	1.2	0.8
ECs2887	*baeR*	hypothetical protein	0.9	1.3
ECs3027	*rffH*	putative salicylate hydroxylase	1.2	0.8

ECs number	Gene[a]	Function	WT/ ΔgadE Exp phase	WT/ ΔgadE Stat phase
ECs3041	*mglA*	ATP-binding component of methyl-galactoside transport and galactose taxis	1.2	0.8
ECs3174	-	putative aminotransferase	1.1	0.8
ECs3213	*aroC*	chorismate synthase	0.8	1.8
ECs3221	-	putative fimbrial usher	1.1	0.8
ECs3223		hypothetical protein	0.8	1.4
ECs3224	-	putative enzyme	0.7	2.4
ECs3228	-	hypothetical protein	1.2	0.7
ECs3306	*amiA*	N-acetylmuramoyl-l-alanine amidase I	1.2	0.8
ECs3307	*hemF*	coproporphyrinogen III oxidase	1.2	0.8
ECs3340	*dapA*	dihydrodipicolinate synthase	1.1	0.8
ECs3389	*pepB*	putative peptidase	1.2	0.8
ECs3390	*yfhJ*	hypothetical protein	1.1	0.8
ECs3391	*fdx*	[2FE-2S] ferredoxin, electron carrer protein	1.1	0.8
ECs3393	*yfhE*	hypothetical protein	1.1	0.8
ECs3395		NifU-like protein	1.2	0.7
ECs3396	*yfhO*	cysteine desulfurase	1.1	0.8
ECs3397	-	hypothetical protein	1.2	0.8
ECs3416	*yphH*	hypothetical protein	1.2	0.8
ECs3422	*yfhK*	putative prophage integrase	0.8	1.6
ECs3540	*proV*	ATP-binding component of transport system for glycine, betaine and proline	0.9	1.2
ECs3595	*rpoS*	RNA polymerase, sigma S (sigma38) factor	0.8	1.5
ECs3596	*nlpD*	lipoprotein	0.9	1.3

ECs number	Gene[a]	Function	WT/ ΔgadE Exp phase	WT/ ΔgadE Stat phase
ECs3696	*lysR*	positive regulator for lys	0.9	1.3
ECs3704	*yqeI*	putative sensory transducer	0.6	2.4
ECs3720	*yraJ*	putative transcriptional regulator	1.1	0.9
ECs3721	*phnE*	putative integral membrane protein-component of type III secretion apparatus	1.1	0.8
ECs3911	*ygiN*	hypothetical protein	0.9	1.3
ECs3929	*ribB*	3,4 dihydroxy-2-butanone-4-phosphate synthase	1.1	0.8
ECs3969	*ygjR*	hypothetical protein	0.9	1.3
ECs4000	*yhaB*	hypothetical protein	0.9	1.4
ECs4188	*hopD*	leader peptidase HopD	1.1	0.8
ECs4213	*yhfG*	hypothetical protein	0.9	1.3
ECs4294	*yhhA*	hypothetical protein	0.8	1.4
ECs4363	*yhiM*	hypothetical protein	0.5	2.8
ECs4389	*hdeB*	member of acid fitness island	0.6	2.8
ECs4390	*hdeA*	member of acid fitness island	0.3	11.1
ECs4391	*hdeD**	member of acid fitness island	0.6	2.5
ECs4392	*gadE*	member of acid fitness island	0.5	4.4
ECs4397	*gadA**	glutamate decarboxylase isozyme	0.4	6.4
ECs4483	*lldD*	L-lactate dehydrogenase	0.5	4.8
ECs4484	*yibK*	hypothetical protein	0.4	5.4
ECs4485	*cysE*	serine acetyltransferase	0.8	1.6
ECs4490	*yibO*	putative 2,3-bisphosphoglycerate-independent phosphoglycerate mutase	0.4	5.9
ECs4491	*yibP*	putative membrane protein	0.6	3.4

ECs number	Gene[a]	Function	WT/ ΔgadE Exp phase	WT/ ΔgadE Stat phase
ECs4579	*yfjI*	hypothetical protein	0.9	1.3
ECs4608	-	unknown function	0.9	1.2
ECs4636	*dnaN*	DNA polymerase III, beta-subunit	0.6	2.8
ECs4683	*gidA*	glucose-inhibited division; chromosome replication?	1.1	0.8
ECs4737	-	hypothetical protein	0.5	5.4
ECs4738	*cyaY*, -	Iron binding frataxin homolog	0.6	2.6
ECs4739	*dapF*	diaminopimelate epimerase	0.6	2.8
ECs4745	-	hypothetical protein	0.8	1.6
ECs4854	*mcrC*	unknown function	0.7	1.9
ECs4891	*udhA*	soluble pyridine nucleotide transhydrogenase	1.1	0.8
ECs5030	*yjbM*	hypothetical protein	1.2	0.8
ECs5080	*phnK*	hypothetical protein	1.2	0.8
ECs5092	*yjcZ*	hypothetical protein	1.1	0.8
ECs5116	*yjdC*	hypothetical protein	0.7	2.5
ECs5124	*mopA*	GroEL, chaperone Hsp60, heat shock protein	1.1	0.8
ECs5163	*aidB*	putative acyl coenzyme A dehydrogenase	0.8	1.5
ECs5215	*nrdD*	anaerobic ribonucleoside-triphosphate reductase	0.9	1.2
ECs5253	*hemN*	partial putative integrase	0.8	1.2
ECs5334	*osmY*	hyperosmotically inducible periplasmic protein	0.8	1.3
ECs5496	-	unknown function	0.9	1.3

[a] The genes marked with asterisks have putative GAD boxes upstream of their sequence.

APPENDIX 4

TABLE 2.S3: Primer sequences used for Q-PCR (SYBR Green chemistry).

Primer	Sequence (5' to 3')	Annealing temperature
16s-982	CGATGCAACGCGAAGAACCT	55°C
16s-1143	CCGGACCGCTGGCAACAAA	
tir-664	ACTTCCAGCCTTCGTTCAGA	57°C
tir-869	TTCTGGAACGCTTCTTTCGT	
espZ-46	GCGACCTCACTCAGTGGAA	55°C
espZ-193	CCGCTGCAATACCTGTACCT	
espD-F	GGTTACAAGTCGCACTGAGGA	59°C
espD-R	CCAGGGATAACAGAGTGACCA	
espF-F	AGCAGCCAGGTGACTTCATT	54°C
espF-R	CTGTGCAATGGGCGGTAAAG	
eae-2188	GCCGGTAAAGCGACTGTTAG	55°C
eae-2325	ATTAGGCAACTCGCCTCTGA	
espA-128	AGGCTGCGATTCTCATGTTT	57°C
espA-310	GAAGTTTGGCTTTCGCATTC	
espB-319	TCAGCATTGGGGATCTTAGG	57°C
espB-487	CTGCGACATCAGCAACACTT	
gadC-571	TGCAAGACCTTCTTCCCTGA	55°C
gadC-694	GCCCTGGGTTACTCATTTCA	
hdeA-F	GAAGATTTCCTGGCTGTGGA	59°C
hdeA-R	ACGGTTGCAATACCCTGAAC	
hdeB-F	CACTGGTGAACGCACAATCT	59°C
hdeB-R	TTTCTTCATGCAGCATCCAC	
gadX-F	TTACAACCGAACATGCGAAC	59°C
gadX-R	CAGACTTGGACTCATCAACAGC	
evgA-F	GAGTTGACTGAAGGCGGAAG	59°C
evgA-R	GGTCATTTTTAGCGGAGACG	

APPENDIX 5

TABLE 4.S1: Genes significantly differentially expressed between clinical genotype 1 and bovine-biased genotype 5 as detected by MAANOVA/Fs test (FDR < 0.1)

ECs no.	Gene	Function	C/B
Genes up-regulated in clinical genotype 1			
ECs0124	*speD*	S-adenosylmethionine decarboxylase	1.72
ECs0125	*speE*	spermidine synthase	1.78
ECs0228		hypothetical protein	2.39
ECs0239		hypothetical protein	2.09
ECs0273		hypothetical protein	4.48
ECs0274		unknown	1.91
ECs0288		hypothetical protein	2.07
ECs0290		hypothetical protein	1.76
ECs0291		hypothetical protein	1.71
ECs0299		putative DNA binding protein	1.71
ECs0331		putative NADH-dependent flavin oxidoreductase	1.51
ECs0445	*rdgC*	recombination associated protein	1.60
ECs0603		Rhs core protein	1.70
ECs0744		putative fimbrial-like protein	3.11
ECs0754	*sucD*	succinyl-CoA synthetase alpha subunit	1.59
ECs0764		putative glutamate mutase S - hypothetical protein	1.65
ECs0815		anti-termination protein	1.53
ECs0829		putative protease/scaffold protein	2.13
ECs0837		putative tail length tape measure protein precursor	1.52
ECs0852	*bioA*	7,8-diaminopelargonic acid synthetase	1.60
ECs0853	*bioB*	biotin synthetase	1.53
ECs1053	*yccK*	putative sulfite reductase	1.58
ECs1127		hypothetical protein	3.13
ECs1139	*yccZ*	hypothetical protein	1.53
ECs1166		hypothetical protein	1.95
ECs1167		hypothetical protein	2.15
ECs1220		putative terminase large subunit	1.97
ECs1222		hypothetical protein	2.91
ECs1224		hypothetical protein	1.76
ECs1225		hypothetical protein	1.55
ECs1226		hypothetical protein	1.57
ECs1228		putative tail fiber protein	4.50
ECs1230		hypothetical protein	2.53
ECs1235		hypothetical protein	1.66
ECs1327	*ureG*		2.60
ECs1349		hypothetical protein	3.03
ECs1370		putative glucosyl-transferase	1.92
ECs1378		hypothetical protein	1.79
ECs1386		hypothetical protein	1.68

ECs1386		hypothetical protein	1.83
ECs1389		hypothetical protein	1.78
ECs1814		hypothetical protein	3.00
ECs2116		putative ATP-binding component of a transport system	2.04
ECs2156		hypothetical protein	1.55
ECs2276		putative replication protein	1.76
ECs2332		hypothetical protein	1.73
ECs2813	*sbcB*	exonuclease I	1.97
ECs2947		putative minor tail protein	1.67
ECs3075	*rsuA*	16S pseudouridylate 516 synthase	1.74
ECs3078	*yejK*	nucleoid-associated protein	1.59
ECs3550	*gshA*	gamma-glutamate-cysteine ligase	1.85
ECs4426		putative fimbrial protein precursor	2.17
ECs4565	*sepQ*	TTSS	1.61
ECs4576	*cesD*	Secretion of EspD	1.77
ECs4591	*rorf1*	unknown	1.58
ECs4653		hypothetical protein	2.92
ECs5073		putative ATP-binding component of sugar ABC transporter	1.81
ECs5318	*yjjN*	putative oxidoreductase	1.55
pO157p02	*etpC*	T2SS	2.75
pO157p05	*etpF*	T2SS	1.86
pO157p07	*etpH*	T2SS	1.91
pO157p09	*etpJ*	T2SS	2.01
pO157p10	*etpK*	T2SS	2.22
pO157p11	*etpL*	T2SS	1.75
pO157p12	*etpM*	T2SS	1.92
pO157p14	*etpO*	T2SS	1.69
pO157p18	*hlyA*	hemolysin A	1.73
pO157p24	*repFIB*	RepFIB	1.67
pO157p58	*toxB*	toxin B	1.59
pO157p79		hypothetical protein	2.44
pO157p81		hypothetical protein	2.12

ECs no.	Gene	Function	B/C
Genes up-regulated in bovine-biased genotype 5			
ECs0068	*araC*	transcriptional regulator for ara operon	1.79
ECs0075	*leuD*	isopropylmalate isomerase subunit	2.49
ECs0076	*leuC*	3-isopropylmalate isomerase (dehydratase) subunit	2.45
ECs0077	*leuB*	3-isopropylmalate dehydrogenase	2.40
ECs0231		unknown function	1.62
ECs0419	*tauA*	taurine transport system periplasmic protein	1.65
ECs0420	*tauB*	taurine ATP-binding component of a transport system	1.74
ECs0499	*ybaE*	hypothetical protein	1.67
ECs0538	*ybaS*	putative glutaminase	3.00
ECs0645	*ahpF*	hypothetical protein	1.51
ECs0732		hypothetical protein	2.04
ECs0848		unknown protein encoded by prophage CP-933K	4.20
ECs0868		hypothetical protein	1.86
ECs0890	*dps*	DNA protection during starvation	3.56
ECs0892	*ompX*	outer membrane protein X	1.71

ECs0957		pyruvate oxidase	2.48
ECs0966	*cspD*	cold shock protein	2.49
ECs0992	*ycaL*	putative heat shock protein	1.59
ECs1077		hypothetical protein	1.70
ECs1159		hypothetical protein	2.96
ECs1181		putative anti-termination protein N	3.06
ECs1182		hypothetical protein	3.55
ECs1197	*ninE*	hypothetical protein	2.31
ECs1201	*ninG*	unknown protein encoded by bacteriophage BP-933W	1.74
ECs1311		unknown in IS600	2.02
ECs1428		hypothetical protein	1.63
ECs1508		unknown protein encoded by prophage CP-933N	2.02
ECs1509		unknown protein encoded by prophage CP-933N	2.44
ECs1546		unknown protein encoded by prophage CP-933N	1.64
ECs1683	*ycgB*	putative sporulation protein	2.96
ECs1722	*chaB*	cation transport regulator	2.37
ECs1743	*oppA*	oligopeptide transport; periplasmic binding protein	2.10
ECs1744	*oppB*	oligopeptide transport permease protein	1.56
ECs1746	*oppD*	ATP-binding protein of oligopeptide ABC transport system	1.82
ECs1747	*oppF*	ATP-binding protein of oligopeptide ABC transport system	1.73
ECs1769		fimbrial minor pilin protein precursor	2.60
ECs1769		putative fimbrial minor pilin protein precursor	3.09
ECs1800		putative tail component encoded by cryptic prophage CP-933M	1.53
ECs1831	*yciG*	hypothetical protein	1.55
ECs1878	*ordL*	probable oxidoreductase	1.50
ECs1975		unknown protein encoded by cryptic prophage CP-933M	2.02
ECs2044		putative transport protein	1.54
ECs2053	*yncB*	putative oxidoreductase	1.55
ECs2086	*osmC*	osmotically inducible protein	2.14
ECs2097	*gadC*	GABA-glutamate antiporter	2.48
ECs2098	*gadB*	glutamate decarboxylase isozyme	4.35
ECs2196		hypothetical protein	2.32
ECs2206		unknown protein encoded within prophage CP-933O	2.05
ECs2236		putative tail assembly protein of cryptic prophage CP-933P	1.54
ECs2300	*mlc*	putative NAGC-like transcriptional regulator	1.51
ECs2370	*cfa*	cyclopropane fatty acyl phospholipid synthase	1.96
ECs2384	*lpp*	murein lipoprotein	1.99
ECs2430		hypothetical protein	1.83
ECs2431		hypothetical protein	2.47
ECs2546		hypothetical protein	1.69
ECs2595	*tap*	methyl-accepting chemotaxis protein IV	1.56
ECs2692		hypothetical protein	2.52
ECs2758		hypothetical protein	2.05
ECs2761		unknown protein encoded within prophage CP-933U	2.57
ECs2825	*hisA*	imidazolecarboxamide isomerase	2.14
ECs2900		hypothetical protein	1.50
ECs2942		hypothetical protein	1.74
ECs3018	*yehZ*	putative transport system permease protein	1.73
ECs3061	*fruB*	PTS system, fructose-specific IIA/fpr component	4.47

ECs3092		cytochrome c-type protein	2.19
ECs3104	*ompC*	outer membrane protein 1b	1.51
ECs3111		hypothetical protein	2.52
ECs3189		UDP-galactopyranose mutase	2.24
ECs3194	*argT*	lysine-, arginine-, ornithine-binding periplasmic protein	1.54
ECs3206		putative transport protein	2.43
ECs3326	*talA*	transaldolase A	1.97
ECs3327		transketolase 2 isozyme	2.19
ECs3409		hypothetical protein	1.69
ECs3533		hypothetical protein	2.58
ECs3680	*ppdC*	hypothetical protein	2.08
ECs3839	*mltC*	transport of nucleosides, permease protein	1.57
ECs3973	*uxaA*	altronate hydrolase	1.52
ECs3974	*uxaC*	uronate isomerase	1.57
ECs3980	*yqjD*	hypothetical protein	1.97
ECs3981		hypothetical protein	1.91
ECs4212		induced in stationary phase, recognized by rpoS	2.37
ECs4213		hypothetical protein	2.21
ECs4273	*glgP*	glycogen phosphorylase	1.52
ECs4275	*glgC*	glucose-1-phosphate adenylyltransferase	1.50
ECs4276	*glgX*	part of glycogen operon, a glycosyl hydrolase	1.55
ECs4278	*asd*	aspartate-semialdehyde dehydrogenase	1.61
ECs4377	*slp*	outer membrane protein	2.45
ECs4389	*hdeB*	periplasmic chaperone	3.37
ECs4390	*hdeA*	protection from organic acid metabolites	5.60
ECs4391	*hdeD*	acid resistance at high cell density	2.67
ECs4394	*yhiV*	multidrug efflux pump protein	2.24
ECs4395	*gadW*	ARAC-type GAD system regulator	1.79
ECs4397	*gadA*	glutamate decarboxylase isozyme	1.87
ECs4412		hypothetical protein	1.77
ECs4427		hypothetical protein	1.74
ECs4440		hypothetical protein	1.82
ECs4483	*lldD*	L-lactate dehydrogenase	1.70
ECs4490	*yibO*	Putative phosphoglycerate mutase	1.56
ECs4660	*phoU*	negative regulator for pho regulon	1.78
ECs4706	*ilvA*	threonine deaminase (dehydratase)	1.80
ECs4734	*hemD*	uroporphyrinogen III synthase	2.13
ECs4737		hypothetical protein	3.38
ECs4760	*ysgA*	putative enzyme	1.81
ECs4871	*katG*	catalase; hydroperoxidase HPI(I)	1.71
ECs4957		hypothetical protein	2.37
ECs4958		hypothetical protein	2.89
ECs4959		hypothetical protein	2.17
ECs4969		putative portal protein	1.90
ECs4970		hypothetical protein	2.61
ECs5038		hypothetical protein	2.21
ECs5039		hypothetical protein	2.18
ECs5120		aspartate ammonia-lyase (aspartase)	2.83
ECs5133	*frdC*	fumarate reductase, anaerobic	1.54
ECs5253		partial putative integrase	1.87

ECs5303		hypothetical protein	1.84
ECs5312		hypothetical protein	2.52
ECs5313	*yjiY*	putative carbon starvation protein	3.58
ECs5325	*yjjQ*	putative regulator	1.67
ECs5326	*bglJ*	2-component transcriptional regulator	2.13
ECs5334	*osmY*	hyperosmotically inducible periplasmic protein	3.01

APPENDIX 6

TABLE 4.S2: Genes differentially expressed between clinical genotype 1 and bovine-biased genotype 5 as detected by SAM (FDR < 0.05)

ECs no.	Gene	Function	C/B
Genes upregulated in clinical genotype 1			
ECs0026	*rpsT*	30S ribosomal subunit protein S20	2.04
ECs0063	*rapA*	probable ATP-dependent RNA helicase	2.27
ECs0201	*abc*	ATP-binding component of a transporter	2.29
ECs0207	*dniR*	transcriptional regulator for nitrite reductase	3.17
ECs0228		hypothetical protein	2.27
ECs0239		hypothetical protein	2.15
ECs0251	*yafK*	hypothetical protein	1.99
ECs0265	*gpt*	guanine-hypoxanthine phosphoribosyltransferase	3.14
ECs0273		unknown protein from prophage CP-933H	4.14
ECs0274		repressor protein CI	1.89
ECs0288		unknown protein encoded by IS2	1.92
ECs0311		putative transcriptional regulator	2.17
ECs0338		putative dehydrogenase	1.89
ECs0351		Unknown function	1.96
ECs0456	*queA*	synthesis of queuine in tRNA	2.19
ECs0477	*yajK*	putative oxidoreductase	1.90
ECs0488	*yajG*	putative polymerase/proteinase	1.88
ECs0510	*ylaB*	hypothetical protein	2.17
ECs0522	*apt*	adenine phosphoribosyltransferase	2.12
ECs0584	*purK*	phosphoribosylaminoimidazole carboxylase	1.84
ECs0586	*ybbF*	hypothetical protein	1.83
ECs0588	*cysS*	cysteine tRNA synthetase	1.88
ECs0673	*mrdA*	cell elongation	2.18
ECs0699	*yleA*	hypothetical protein	1.92
ECs0716	*ybfE*	hypothetical protein	2.06
ECs0744	*ybgD*	putative fimbriae structural protein	3.34
ECs0771	*ybgC*	hypothetical protein	2.19
ECs0772	*tolQ*	inner membrane protein	2.17
ECs0800	*ybhC*	putative integrase encoded by prophage CP-933K	2.11
ECs0829		putative protease encoded in prophage CP-933K	2.11
ECs0883		putative outer membrane receptor for iron transport	2.07
ECs0896	*ybiS*	hypothetical protein	2.26
ECs0915	*yliG*	hypothetical protein	1.97
ECs0939	*ybjF*	putative enzyme	1.85
ECs0944	*artM*	arginine 3rd transport system permease protein	1.53
ECs0945	*artQ*	arginine 3rd transport system permease protein	1.81
ECs0946	*artI*	arginine 3rd transport system periplasmic binding protein	1.94
ECs0947	*artP*	ATP-binding component of 3rd arginine transport	1.65

		system	
ECs0960	*ybjE*	putative surface protein	2.34
ECs0997	*msbA*	ATP-binding transport protein	1.94
ECs1127		unknown protein encoded by cryptic prophage CP-933M	2.92
ECs1142	*ymcC*	putative regulator	1.81
ECs1166		hypothetical protein	2.01
ECs1167		hypothetical protein	2.09
ECs1192		unknown protein encoded by bacteriophage BP-933W	1.78
ECs1203		antitermination protein Q of bacteriophage BP-933W	2.60
ECs1220		partial putative terminase large subunit of bacteriophage BP-933W	1.86
ECs1222		unknown protein encoded by bacteriophage BP-933W	2.67
ECs1228		putative tail fiber protein of bacteriophage BP-933W	3.90
ECs1230		unknown protein encoded by bacteriophage BP-933W	2.54
ECs1233		unknown protein encoded by bacteriophage BP-933W	2.36
ECs1234		putative outer membrane protein	1.73
ECs1236		putative outer membrane protein of bacteriophage BP-933W	2.00
ECs1327	*ureG*	Unknown function	2.51
ECs1349		Unknown function	2.97
ECs1362		Unknown function	2.05
ECs1370		putative glucosyltransferase	1.93
ECs1378		Unknown function	1.67
ECs1379		hypothetical protein	1.79
ECs1386		Unknown function	1.79
ECs1387		Unknown function	1.86
ECs1398		Unknown function	1.80
ECs1433	*yceA*	hypothetical protein	1.81
ECs1464	*yceC*	hypothetical protein	2.00
ECs1494	*ycfU*	hypothetical protein	1.77
ECs1500	*potC*	spermidine/putrescine transport system permease	1.87
ECs1564		IS2 hypothetical protein	2.16
ECs1605		hypothetical protein	1.94
ECs1612		putative replication protein P of bacteriophage BP-933W	1.74
ECs1646		putative tail component of prophage CP-933K	2.82
ECs1703	*ycgC*	putative PTS system enzyme I	2.04
ECs1705		putative dihydroxyacetone kinase (EC 2.7.1.2)	1.69
ECs1709		peptidyl-tRNA hydrolase	1.68
ECs1733		a protaminelike protein	1.94
ECs1740	*tdk*	thymidine kinase	1.95
ECs1814		hypothetical protein	2.90
ECs1854	*pyrF*	orotidine-5'-phosphate decarboxylase	2.13
ECs1859	*rnb*	RNase II, mRNA degradation	1.87

ECs1872	*ymjA*	hypothetical protein	1.92
ECs1928	*ydaO*	hypothetical protein	1.86
ECs2021	*aldA*	aldehyde dehydrogenase, NAD-linked	2.31
ECs2099	*pqqL*	putative peptidase	2.51
ECs2100	*yddB*	hypothetical protein	2.13
ECs2116		putative ATP-binding component of a transport system	2.08
ECs2229		unknown protein encoded by prophage CP-933O	1.79
ECs2259		putative lysozyme	2.83
ECs2337		hypothetical protein	1.83
ECs2367	*purR*	transcriptional repressor for pur regulon, glyA, glnB, prsA, speA	2.19
ECs2516	*yeaZ*	hypothetical protein	1.89
ECs2653	*yecF*	hypothetical protein	2.42
ECs2813	*sbcB*	exonuclease I, 3' --> 5' specific	2.01
ECs2816	*yeeF*	putative amino acid/amine transport protein	3.83
ECs2828	*wzzB*	regulator of length of O-antigen component of lipopolysaccharide chains	2.39
ECs2873	*udk*	uridine/cytidine kinase	2.08
ECs2889	*yegQ*	hypothetical protein	2.46
ECs2973	*stx1b*	shiga-like toxin 1 subunit B encoded within prophage CP-933V	19.82
ECs2974	*stx1a*	shiga-like toxin 1 subunit A encoded within prophage CP-933V	15.10
ECs2975		putative antitermination protein Q for prophage CP-933V	1.89
ECs2992		unknown protein encoded within prophage CP-933V	4.28
ECs3002		putative exonuclease of bacteriophage BP-933W	1.76
ECs3048	*lysP*	lysine-specific permease	2.17
ECs3067	*spr*	putative lipoprotein	1.63
ECs3079	*yejL*	hypothetical protein	1.82
ECs3223		hypothetical protein	2.10
ECs3309		hypothetical protein	1.94
ECs3362	*purN*	phosphoribosylglycinamide formyltransferase 1	1.84
ECs3380	*ndk*	nucleoside diphosphate kinase	2.14
ECs3399	*suhB*	extragenic suppressor	3.30
ECs3530	*stpA*	DNA-binding protein	2.24
ECs3550		unknown protein encoded by bacteriophage BP-933W	1.84
ECs3620	*ygcM*	putative 6-pyruvoyl tetrahydrobiopterin synthase	1.80
ECs3673	*mltA*	membrane-bound lytic murein transglycosylase A	2.05
ECs3689		putative transport protein	2.18
ECs3818	*metK*	methionine adenosyltransferase 1	1.91
ECs3948		30S ribosomal subunit protein S21	2.27
ECs4043	*deaD*	inducible ATP-independent RNA helicase	2.07
ECs4054	*secG*	protein export - membrane protein	2.17
ECs4069	*yrbA*	hypothetical protein	1.72
ECs4132	*yhdG*	putative dehydrogenase	2.50
ECs4133	*fis*	site-specific DNA inversion stimulation factor	2.18
ECs4214	*ppiA*	peptidyl-prolyl cis-trans isomerase A (rotamase A)	1.83
ECs4238		peptidoglycan synthetase; penicillin-binding	1.96

		protein 1A	
ECs4426		putative fimbrial subunit	1.89
ECs4461	*yiaU*	putative transcriptional regulator LYSR-type	2.35
ECs4550	*espF*	secreted protein/effector	1.97
ECs4552	*escF*	T3SS	2.09
ECs4566	*orf16*	secretion of translocators	2.16
ECs4570	*orf12*	T3SS	2.13
ECs4574	*sepD*		2.04
ECs4579	*rorf3*		1.91
ECs4581	*escT*	T3SS	1.98
ECs4583	*escR*	T3SS	1.91
ECs4590	*espG*	secreted protein/effector	1.88
ECs4606	*uhpA*	response regulator, positive activator of uhpT transcription	2.24
ECs4638		50S ribosomal subunit protein L34	2.38
ECs4639	*rnpA*	RNase P, protein component; protein C5; processes tRNA, 4.5S RNA	2.73
ECs4653		Unknown function	2.97
ECs4976		hypothetical protein	1.83
ECs4982		hypothetical protein	1.84
ECs5074		Unknown function	1.80
ECs5147	*miaA*	hypothetical protein	2.09
ECs5246		Unknown function	2.59
ECs5252		putative transcriptional regulator	1.88
ECs5320	*yjjA*	putative glycoprotein/receptor	1.72
pO157p02	*etpC*	T2SS	3.04
pO157p05	*etpF*	T2SS	1.98
pO157p09	*etpJ*	T2SS	2.21
pO157p10	*etpK*	T2SS	2.33
pO157p12	*etpM*	T2SS	2.02
pO157p79		hypothetical protein	2.43
pO157p80		hypothetical protein	2.07
pO157p81		hypothetical protein	2.30
pOSAK1_02		hypothetical protein	8.58
pOSAK1_03		hypothetical protein	20.66

ECs no.	Gene	Function	B/C
Genes upregulated in bovine-biased genotype 5			
ECs0007	*yaaJ*	inner membrane transport protein	2.01
ECs0068	*araC*	transcriptional regulator for ara operon	1.84
ECs0075	*leuD*	isopropylmalate isomerase subunit	2.77
ECs0076	*leuC*	3-isopropylmalate isomerase subunit	2.68
ECs0077	*leuB*	3-isopropylmalate dehydrogenase	2.62
ECs0078	*leuA*	2-isopropylmalate synthase	1.94
ECs0118	*aceE*	pyruvate dehydrogenase (decarboxylase component)	3.10
ECs0119	*aceF*	pyruvate dehydrogenase (dihydrolipoyltransacetylase component)	3.61
ECs0169	*glnD*	uridylyltransferase acts on regulator of glnA	1.83
ECs0191	*yaeO*	hypothetical protein	2.19
ECs0287	*yhiV*	putative AraC-type regulatory protein encoded in	1.99

		prophage CP-933H	
ECs0357	*betA*	choline dehydrogenase, a flavoprotein	2.02
ECs0358	*betB*	NAD+-dependent betaine aldehyde dehydrogenase	1.70
ECs0408	*mhpT*	putative transport protein	2.19
ECs0504	*glnK*	nitrogen regulatory protein P-II 2	11.58
ECs0505	*amtB*	probable ammonium transporter	14.94
ECs0537	*ybaR*	putative ATPase	1.98
ECs0538	*ybaS*	putative glutaminase	3.30
ECs0646	*ybdQ*	hypothetical protein	2.50
ECs0681	*ybeL*	putative alpha helical protein	2.24
ECs0692	*gltK*	glutamate/aspartate transport system permease	1.91
ECs0693	*gltJ*	glutamate/aspartate transport system permease	2.16
ECs0694	*ybeJ*	putative periplasmic binding transport protein	3.52
ECs0732	*ybgA*	hypothetical protein	2.00
ECs0780	*ybgR*	putative transport system permease protein	3.00
ECs0781		putative homeobox protein	2.07
ECs0848		unknown protein encoded by prophage CP-933K	4.00
ECs0868	*ybhP*	hypothetical protein	1.88
ECs0873		putative membrane protein	1.77
ECs0889	*glnH*	permease of periplasmic glutamine-binding protein	2.12
ECs0890	*dps*	global regulator, starvation conditions	3.98
ECs0892	*ompX*	outer membrane protein X	1.70
ECs0916		putative receptor	9.18
ECs0927	*yjiU*	hypothetical protein	1.99
ECs0957	*poxB*	pyruvate oxidase	2.75
ECs0966	*cspD*	cold shock protein	2.55
ECs0968	*clpA*	ATP-binding component of serine protease	1.99
ECs1002		hypothetical protein	4.08
ECs1037	*rmf*	ribosome modulation factor	3.11
ECs1077	*gef*	hypothetical protein	1.71
ECs1091		putative transcriptional regulator	2.22
ECs1159	*yccJ*	hypothetical protein	3.47
ECs1159	*wrbA*	trp repressor binding protein	2.94
ECs1180	*polB*	hypothetical protein	2.15
ECs1181	*ycfX*	putative anti-termination protein N	3.37
ECs1182	*citB*	hypothetical protein	3.79
ECs1183	*yjdK*	hypothetical protein	2.26
ECs1184	*flgK*	hypothetical protein	2.19
ECs1185	*yiaW*	putative cI repressor protein	2.49
ECs1197	*ycfO*	hypothetical protein	2.31
ECs1201		unknown protein encoded by bacteriophage BP-933W	1.77
ECs1252	*ycdG*	putative transport protein	3.56
ECs1253		hypothetical protein	2.61
ECs1254		putative enzyme	6.83
ECs1255		putative acetyltransferase	10.01
ECs1256		hypothetical protein	8.93
ECs1257		putative synthetase	5.59
ECs1311		unknown in IS600	2.03

ECs1316		putative diacylglycerol kinase	2.06
ECs1360		putative sensor-type regulator	1.92
ECs1442	*grxB*	glutaredoxin 2	2.08
ECs1508	*yjiO*	unknown protein encoded by prophage CP-933N	2.01
ECs1509	*rpoC*	unknown protein encoded by prophage CP-933N	2.55
ECs1572	*pepT*	putative peptidase T	2.66
ECs1683	*ycgB*	putative sporulation protein	3.23
ECs1684	*dadA*	D-amino acid dehydrogenase subunit	1.94
ECs1710	*ychH*	hypothetical protein	4.95
ECs1722	*chaB*	cation transport regulator	2.45
ECs1723	*chaC*	cation transport regulator	2.23
ECs1728	*narK*	nitrite extrusion protein	1.91
ECs1741	*adhE*	CoA-linked acetaldehyde dehydrogenase	3.83
ECs1743	*oppA*	oligopeptide transport	2.18
ECs1746	*oppD*	oligopeptide transport	1.83
ECs1747	*oppF*	oligopeptide transport	1.77
ECs1769		putative phage replication protein	3.17
ECs1775		hypothetical protein	2.12
ECs1914	*ydaA*	hypothetical protein	2.72
ECs1975	*adhE*	unknown protein encoded by cryptic prophage CP-933M	2.04
ECs2084	*rpsV*	30S ribosomal subunit protein S22	2.63
ECs2086	*osmC*	osmotically inducible protein	2.27
ECs2087		putative ATP-binding component of a transport system	5.17
ECs2088		putative ATP-binding component of a transport system	6.60
ECs2089		putative transport protein	2.75
ECs2090		putative transport system permease protein	3.55
ECs2091		putative hemin-binding lipoprotein	8.57
ECs2092		hypothetical protein	9.17
ECs2097	*gadC*	GABA-gluatame antiporter	2.50
ECs2098	*gadB*	glutamate decarboxylase isozyme	4.70
ECs2196	*mcrC*	hypothetical protein	2.42
ECs2203	*pstC*	hypothetical protein	2.14
ECs2206	*ydaU*	unknown protein encoded within prophage CP-933O	2.13
ECs2209	*ybjT*	putative repressor protein encoded within prophage CP-933O	1.71
ECs2212	*ygbD*	hypothetical protein	2.11
ECs2299		hypothetical protein	2.91
ECs2370	*cfa*	cyclopropane fatty acyl phospholipid synthase	2.10
ECs2384	*lpp*	murein lipoprotein	1.95
ECs2429	*pfkB*	6-phosphofructokinase II; suppressor of pfkA	1.90
ECs2430		hypothetical protein	1.95
ECs2431		hypothetical protein	2.69
ECs2450	*ydjS*	hypothetical protein	5.68
ECs2451		hypothetical protein	6.96
ECs2452		putative aldehyde dehydrogenase	12.07
ECs2454	*cstC*	acetylornithine delta-aminotransferase	9.53
ECs2467	*gdhA*	NADP-specific glutamate dehydrogenase	2.12

ECs2492	*yeaG*	hypothetical protein	6.18
ECs2543		hypothetical protein	2.46
ECs2603	*yecG*	putative regulator	2.78
ECs2604	*otsA*	trehalose-6-phosphate synthase	2.42
ECs2614	*yecH*	hypothetical protein	2.10
ECs2670	*yedL*	hypothetical protein	2.05
ECs2692		hypothetical protein	2.70
ECs2705	*yedU*	hypothetical protein	1.95
ECs2712		putative cytochrome	1.80
ECs2737		putative transcriptional regulator	2.29
ECs2737	*uxaB*	putative transcriptional regulator	2.32
ECs2758	*araF*	hypothetical protein	2.13
ECs2761	*aas*	unknown protein encoded within prophage CP-933U	2.61
ECs2783	*cbl*	transcriptional regulator cys regulon	9.76
ECs2784	*nac*	nitrogen assimilation control protein	12.69
ECs2792		hypothetical protein	2.28
ECs2825	*hisA*	imidazolecarboxamide isomerase	1.90
ECs2882		putative membrane protein	1.87
ECs2887	*baeR*	hypothetical protein	2.54
ECs2927	*yehL*	hypothetical protein	2.15
ECs2942	*ymfM*	hypothetical protein	1.71
ECs3018	*yehZ*	putative transport system permease protein	1.68
ECs3035	*cdd*	cytidine/deoxycytidine deaminase	2.79
ECs3058	*yeiC*	putative kinase	3.99
ECs3060	*fruK*	fructose-1-phosphate kinase	2.04
ECs3061	*fruB*	PTS system, fructose-specific IIA/fpr component	5.30
ECs3092		cytochrome c-type protein	2.12
ECs3097	*napF*	ferredoxin-type protein: electron transfer	2.48
ECs3111		hypothetical protein	2.46
ECs3189	*glf*	UDP-galactopyranose mutase	2.11
ECs3206		putative transport protein	2.34
ECs3226		hypothetical protein	3.28
ECs3257		hypothetical protein	2.52
ECs3259		putative aminotransferase	3.09
ECs3326	*talA*	transaldolase A	2.10
ECs3327	*tktB*	transketolase 2 isozyme	2.29
ECs3346	*hyfD*	hydrogenase 4 membrane subunit	2.51
ECs3358		putative DNA replication factor	3.09
ECs3445	*yfiD*	putative formate acetyltransferase	6.10
ECs3460	*yfiA*	putative yhbH sigma 54 modulator	5.01
ECs3519		hypothetical protein	6.97
ECs3522	*gabD*	succinate-semialdehyde dehydrogenase	3.69
ECs3523	*gabT*	4-aminobutyrate aminotransferase activity	3.22
ECs3533		hypothetical protein	2.88
ECs3543	*proX*	hypothetical protein	2.47
ECs3582	*hypA*	pleiotrophic effects on 3 hydrogenase isozymes	2.96
ECs3583	*hypB*	hydrogenase isoenzyme HypB	2.61
ECs3643	*chpR*	suppressor of inhibitory function of ChpA	2.10
ECs3644	*relA*	(p)ppGpp synthetase I (GTP pyrophosphokinase)	2.35
ECs3655	*ygdH*	hypothetical protein	1.85

ECs3669		hypothetical protein	1.85
ECs3680		hypothetical protein	2.08
ECs3689		hypothetical protein	7.67
ECs3799	*ybbF*	hypothetical protein	2.82
ECs3842	*yqgA*	putative transport protein	2.23
ECs3891		hypothetical protein	4.13
ECs3904		putative transport periplasmic protein	2.72
ECs3931	*glgS*	glycogen biosynthesis, rpoS dependent	2.27
ECs3980	*yqjD*	hypothetical protein	2.11
ECs3981	*yqjE*	hypothetical protein	1.95
ECs4034	*yhbO*	hypothetical protein	1.75
ECs4038	*yhbT*	hypothetical protein	2.44
ECs4089	*arcB*	aerobic respiration sensor-response protein	1.99
ECs4091	*gltB*	glutamate synthase, large subunit	1.91
ECs4141		putative periplasmic binding transport protein	5.55
ECs4142		putative transport system permease protein	3.15
ECs4212	*fic*	induced in stationary phase	2.64
ECs4213	*yhfG*	hypothetical protein	2.45
ECs4216	*nirB*	nitrite reductase (NAD(P)H) subunit	1.82
ECs4250	*feoA*	ferrous iron transport protein A	2.68
ECs4309		high-affinity amino acid transport system	4.99
ECs4323		hypothetical protein	7.11
ECs4366	*yhiO*	hypothetical protein	2.32
ECs4367	*uspA*	universal stress protein	2.64
ECs4377	*slp*	outer membrane protein	2.52
ECs4389	*hdeB*	periplasmic chaperone	3.38
ECs4390	*hdeA*	protection from organic acid metabolites	6.12
ECs4391	*hdeD*	acid resistance at high cell density	3.64
ECs4392	*gadE*	central activator of GAD system	2.56
ECs4394	*yhiV*	multidrug efflux pump protein	2.31
ECs4395	*gadW*	ARAC-type GAD system regulator	1.77
ECs4397	*gadA*	glutamate decarboxylase isozyme	1.92
ECs4402	*yhjD*	hypothetical protein	1.88
ECs4412	*yhjN*	hypothetical protein	1.84
ECs4424	*dppA*	dipeptide transport protein	1.89
ECs4427	*yfdO*	hypothetical protein	1.75
ECs4440	*yiaG*	hypothetical protein	1.86
ECs4456	*yiaI*	hypothetical protein	1.95
ECs4477	*mtlR*	repressor for mtl	2.13
ECs4615	*yidF*	putative transcriptional regulator	1.93
ECs4624	*yidP*	putative transcriptional regulator	2.35
ECs4660	*phoU*	negative regulator for pho regulon	1.86
ECs4706	*ilvA*	threonine deaminase (dehydratase)	1.80
ECs4708	*ilvC*	ketol-acid reductoisomerase	3.70
ECs4734	*hemD*	uroporphyrinogen III synthase	2.20
ECs4737		orf, hypothetical protein	3.53
ECs4760	*ysgA*	putative enzyme	1.91
ECs4787	*yihA*	orf; Unknown function	1.98
ECs4790	*glnG*	nitrogen regulator I	2.64
ECs4803	*yihS*	hypothetical protein	2.44

ECs4847	*yiiS*	hypothetical protein	2.08
ECs4848	*yiiT*	putative regulator	2.15
ECs4856	*menG*	menaquinone biosynthesis	1.91
ECs4893	*yheN*	putative citrate permease	1.78
ECs4956	*yraI*	hypothetical protein	1.91
ECs4957		hypothetical protein	2.47
ECs4958	*zipA*	hypothetical protein	2.23
ECs4968	*purH*	hypothetical protein	1.81
ECs4969	*menC*	putative portal protein	1.98
ECs4970	*hsdM*	hypothetical protein	2.77
ECs4973	*gyrB*	putative protease protein	2.11
ECs5012	*yjbH*	hypothetical protein	2.10
ECs5036	*tyrB*	tyrosine aminotransferase	1.92
ECs5038	*yjbQ*	hypothetical protein	2.45
ECs5039	*yjbR*	hypothetical protein	2.34
ECs5052	*nrfA*	periplasmic cytochrome c(552): plays a role in nitrite reduction	2.33
ECs5061	*fdhF*	selenopolypeptide subunit of formate dehydrogenase H	2.08
ECs5100	*melR*	regulator of melibiose operon	2.01
ECs5120		aspartate ammonia-lyase (aspartase)	3.15
ECs5134	*frdB*	fumarate reductase, anaerobic, iron-sulfur protein subunit	2.22
ECs5135	*frdA*	fumarate reductase, anaerobic, flavoprotein subunit	2.49
ECs5184	*ytfB*	hypothetical protein	2.14
ECs5189		hypothetical protein	2.05
ECs5190	*ytfG*	putative oxidoreductase	2.14
ECs5253	*hemN*	partial putative integrase	1.84
ECs5303	*yjiS*	hypothetical protein	1.93
ECs5311	*yjiA*	hypothetical protein	2.09
ECs5312	*yjiX*	hypothetical protein	2.68
ECs5313	*yjiY*	putative carbon starvation protein	3.84
ECs5315	*tsr*	methyl-accepting chemotaxis protein I, serine sensor receptor	1.92
ECs5316	*yjiZ*	putative transport protein, cryptic, orf, joins former yjiZ and yjjL	2.09
ECs5317	*yjjM*	hypothetical protein	2.37
ECs5326	*bglJ*	2-component transcriptional regulator	2.10
ECs5334	*osmY*	hyperosmotically inducible periplasmic protein	3.33

REFERENCES

1. **Abe, H., A. Miyahara, T. Oshima, K. Tashiro, Y. Ogura, S. Kuhara, N. Ogasawara, T. Hayashi, and T. Tobe.** 2008. Global Regulation by Horizontally Transferred Regulators Establishes the Pathogenicity of *Escherichia coli.* DNA Res **15:**25-38.

2. **Abe, H., I. Tatsuno, T. Tobe, A. Okutani, and C. Sasakawa.** 2002. Bicarbonate Ion Stimulates the Expression of Locus of Enterocyte Effacement-Encoded Genes in Enterohemorrhagic *Escherichia coli* O157:H7. Infect. Immun. **70:**3500-3509.

3. **Ake, J. A., S. Jelacic, M. A. Ciol, S. L. Watkins, K. F. Murray, D. L. Christie, E. J. Klein, and P. I. Tarr.** 2005. Relative nephroprotection during *Escherichia coli* O157:H7 infections: association with intravenous volume expansion. Pediatrics **115:**e673-80.

4. **Akerley, B. J., P. A. Cotter, and J. F. Miller.** 1995. Ectopic expression of the flagellar regulon alters development of the *Bordetella*-host interaction. Cell **80:**611-20.

5. **Aldick, T., M. Bielaszewska, W. Zhang, J. Brockmeyer, H. Schmidt, A. W. Friedrich, K. S. Kim, M. A. Schmidt, and H. Karch.** 2007. Hemolysin from Shiga toxin-negative *Escherichia coli* O26 strains injures microvascular endothelium. Microbes Infect **9:**282-90.

6. **Alignan, M., T. Hewezi, M. Petitprez, G. Dechamp-Guillaume, gory, and L. Gentzbittel.** 2006. A cDNA microarray approach to decipher sunflower (*Helianthus annuus*) responses to the necrotrophic fungus *Phoma macdonaldii.* New Phytologist **170:**523-536.

7. **Andreoli, S. P., H. Trachtman, D. W. Acheson, R. L. Siegler, and T. G. Obrig.** 2002. Hemolytic uremic syndrome: epidemiology, pathophysiology, and therapy. Pediatr Nephrol **17:**293-8.

8. **Bergholz, T. M., C. L. Tarr, L. M. Christensen, D. J. Betting, and T. S. Whittam.** 2007. Recent Gene Conversions between Duplicated Glutamate Decarboxylase Genes (*gadA* and *gadB*) in Pathogenic *Escherichia coli.* Mol Biol Evol **24:**2323-2333.

9. **Bergholz, T. M., S. K. Vanaja, and T. S. Whittam.** 2009. Gene expression induced in *Escherichia coli* O157:H7 upon exposure to model apple juice. Appl Environ Microbiol **75:**3542-53.

10. **Bergholz, T. M., and T. S. Whittam.** 2007. Variation in acid resistance among enterohaemorrhagic *Escherichia coli* in a simulated gastric environment. J Appl Microbiol **102:**352-362.

11. **Bergholz, T. M., and T. S. Whittam.** 2007. Variation in acid resistance among enterohaemorrhagic *Escherichia coli* in a simulated gastric environment. Journal of Applied Microbiology **102:**352-362.

12. **Bergholz, T. M., L. M. Wick, W. Qi, J. T. Riordan, L. M. Ouellette, and T. S. Whittam.** 2007. Global transcriptional response of *Escherichia coli* O157:H7 to growth transitions in glucose minimal medium. BMC Microbiol **7:**97.

13. **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:**2217-20.

14. **Besser, T. E., N. Shaikh, N. J. Holt, P. I. Tarr, M. E. Konkel, P. Malik-Kale, C. W. Walsh, T. S. Whittam, and J. L. Bono.** 2007. Greater diversity of Shiga toxin-encoding bacteriophage insertion sites among *Escherichia coli* O157:H7 isolates from cattle than in those from humans. Appl. Environ. Microbiol. **73:**671-679.

15. **Bhagwat, A. A., R. P. Phadke, D. Wheeler, S. Kalantre, M. Gudipati, and M. Bhagwat.** 2003. Computational methods and evaluation of RNA stabilization reagents for genome-wide expression studies. Journal of Microbiological Methods **55:**399-409.

16. **Biase, D. D., A. Tramonti, R. A. John, and F. Bossa.** 1996. Isolation, Overexpression, and Biochemical Characterization of the Two Isoforms of Glutamic Acid Decarboxylase from *Escherichia coli.* Protein Expression and Purification **8:**430-438.

17. **Blattner, F. R., G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277:**1453-74.

18. **Bordi, C., L. Theraulaz, V. Mejean, and C. Jourlin-Castelli.** 2003. Anticipating an alkaline stress through the Tor phosphorelay system in *Escherichia coli.* Mol Microbiol **48:**211-23.

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

20. **Calderwood, S. B., and J. J. Mekalanos.** 1987. Iron regulation of Shiga-like toxin expression in *Escherichia coli* is mediated by the *fur* locus. J Bacteriol **169:**4759-64.

21. **Canales, R. D., Y. Luo, J. C. Willey, B. Austermiller, C. C. Barbacioru, C. Boysen, K. Hunkapiller, R. V. Jensen, C. R. Knight, K. Y. Lee, Y. Ma, B. Maqsodi, A. Papallo, E. H. Peters, K. Poulter, P. L. Ruppel, R. R. Samaha, L. Shi, W. Yang, L. Zhang, and F. M. Goodsaid.** 2006. Evaluation of DNA microarray results with quantitative gene expression platforms. Nat Biotech **24:**1115-1122.

22. **Castanie-Cornet, M.-P., and J. W. Foster.** 2001. Escherichia coli acid resistance: cAMP receptor protein and a 20 bp cis-acting sequence control pH and stationary phase expression of the gadA and gadBC glutamate decarboxylase genes. Microbiology **147:**709-715.

23. **Castanie-Cornet, M.-P., T. A. Penfound, D. Smith, J. F. Elliott, and J. W. Foster.** 1999. Control of acid resistance in *Escherichia coli.* J. Bacteriol. **181:**3525-3535.

24. **Castanie-Cornet, M.-P., H. Treffandier, A. Francez-Charlot, C. Gutierrez, and K. Cam.** 2007. The glutamate-dependent acid resistance system in Escherichia coli: essential and dual role of the His-Asp phosphorelay RcsCDB/AF. Microbiology **153:**238-246.

25. **Chapman, P. A., C. A. Siddons, D. J. Wright, P. Norman, J. Fox, and E. Crick.** 1993. Cattle as a possible source of verocytotoxin-producing *Escherichia coli* O157 infections in man. Epidemiol Infect **111:**439-47.

26. **Chen, J., and M. W. Griffiths.** 1999. Cloning and sequencing of the gene encoding universal stress protein from *Escherichia coli* O157:H7 isolated from Jack-in-a-Box outbreak. Lett Appl Microbiol **29:**103-7.

27. **Chen, J., J. Lozach, E. W. Garcia, B. Barnes, S. Luo, I. Mikoulitch, L. Zhou, G. Schroth, and J.-B. Fan.** 2008. Highly sensitive and specific microRNA expression profiling using BeadArray technology. Nucl. Acids Res. **36:**e87.

28. **Choi, S. H., D. J. Baumler, and C. W. Kaspar.** 2000. Contribution of *dps* to Acid Stress Tolerance and Oxidative Stress Tolerance in *Escherichia coli* O157:H7. Appl. Environ. Microbiol. **66:**3911-3916.

29. **Cody, S. H., M. K. Glynn, J. A. Farrar, K. L. Cairns, P. M. Griffin, J. Kobayashi, M. Fyfe, R. Hoffman, A. S. King, J. H. Lewis, B. Swaminathan, R. G. Bryant, and D. J. Vugia.** 1999. An outbreak of *Escherichia coli* O157:H7 infection from unpasteurized commercial apple juice. Ann Intern Med **130:**202-9.

30. **Cui, X., and G. A. Churchill.** 2003. Statistical tests for differential expression in cDNA microarray experiments. Genome Biol **4:**210.

31. **Cui, X., J. T. Hwang, J. Qiu, N. J. Blades, and G. A. Churchill.** 2005. Improved statistical tests for differential gene expression by shrinking variance components estimates. Biostatistics **6:**59-75.

32. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences **97:**6640-6645.

33. **Deng, W., J. L. Puente, S. Gruenheid, Y. Li, B. A. Vallance, A. Vazquez, J. Barba, J. A. Ibarra, P. O'Donnell, P. Metalnikov, K. Ashman, S. Lee, D. Goode, T. Pawson, and B. B. Finlay.** 2004. Dissecting virulence: Systematic and functional analyses of a pathogenicity island. Proceedings of the National Academy of Sciences **101:**3597-3602.

34. **Diez-Gonzalez, F., T. R. Callaway, M. G. Kizoulis, and J. B. Russell.** 1998. Grain feeding and the dissemination of acid-resistant *Escherichia coli* from cattle. Science **281:**1666-8.

35. **Eaton, K. A., D. I. Friedman, G. J. Francis, J. S. Tyler, V. B. Young, J. Haeger, G. Abu-Ali, and T. S. Whittam.** 2008. Pathogenesis of renal disease due to enterohemorrhagic *Escherichia coli* in germ-free mice. Infect Immun **76:**3054-63.

36. **Elder, R. O., J. E. Keen, G. R. Siragusa, G. A. Barkocy-Gallagher, M. Koohmaraie, and W. W. Laegreid.** 2000. Correlation of enterohemorrhagic *Escherichia coli* O157 prevalence in feces, hides, and carcasses of beef cattle during processing. Proc Natl Acad Sci U S A **97:**2999-3003.

37. **Fann, M. K., and D. O'Rourke.** 2001. Normal bacterial flora of the rabbit gastrointestinal tract: a clinical approach. Seminars in Avian and Exotic Pet Medicine **10:**45-47.

38. **Feng, P., K. A. Lampel, H. Karch, and T. S. Whittam.** 1998. Genotypic and phenotypic changes in the emergence of *Escherichia coli* O157:H7. J Infect Dis **177:**1750-3.

39. **Foster, J. W.** 2004. *Escherichia coli* acid resistance: tales of an amateur acidophile. Nat Rev Micro **2:**898-907.

40. **Fox, J. T., X. Shi, and T. G. Nagaraja.** 2008. *Escherichia coli* O157 in the rectoanal mucosal region of cattle. Foodborne Pathog Dis **5:**69-77.

41. **Garmendia, J., G. Frankel, and V. F. Crepin.** 2005. Enteropathogenic and Enterohemorrhagic *Escherichia coli* Infections: Translocation, Translocation, Translocation. Infect. Immun. **73:**2573-2585.

42. **Gong, S., Z. Ma, and J. W. Foster.** 2004. The Era-like GTPase TrmE conditionally activates *gadE* and glutamate-dependent acid resistance in *Escherichia coli.* Molecular Microbiology **54:**948-961.

43. **Gong, S., H. Richard, and J. W. Foster.** 2003. YjdE (AdiC) Is the Arginine:Agmatine Antiporter Essential for Arginine-Dependent Acid Resistance in *Escherichia coli.* J. Bacteriol. **185:**4402-4409.

44. **Grys, T. E., M. B. Siegel, W. W. Lathem, and R. A. Welch.** 2005. The StcE protease contributes to intimate adherence of enterohemorrhagic *Escherichia coli* O157:H7 to host cells. Infect Immun **73:**1295-303.

45. **Hancock, D., T. Besser, J. Lejeune, M. Davis, and D. Rice.** 2001. The control of VTEC in the animal reservoir. Int J Food Microbiol **66:**71-8.

46. **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.

47. **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 *Eschelichia coli* O157:H7 and Genomic Comparison with a Laboratory Strain K-12. DNA Res **8:**11-22.

48. **Herold, S., J. Siebert, A. Huber, and H. Schmidt.** 2005. Global expression of prophage genes in *Escherichia coli* O157:H7 strain EDL933 in response to norfloxacin. Antimicrob Agents Chemother **49:**931-44.

49. **Herriott, D. E., D. D. Hancock, E. D. Ebel, L. V. Carpenter, D. H. Rice, and T. E. Besser.** 1998. Association of herd management factors with colonization of dairy cattle by Shiga toxin-positive *Escherichia coli* O157. J Food Prot **61**:802-7.

50. **Heuveling, J., A. Possling, and R. Hengge.** 2008. A role for Lon protease in the control of the acid resistance genes of *Escherichia coli.* Mol Microbiol **69**:534-47.

51. **Heuvelink, A. E., F. L. van den Biggelaar, J. Zwartkruis-Nahuis, R. G. Herbes, R. Huyben, N. Nagelkerke, W. J. Melchers, L. A. Monnens, and E. de Boer.** 1998. Occurrence of verocytotoxin-producing *Escherichia coli* O157 on Dutch dairy farms. J Clin Microbiol **36**:3480-7.

52. **Hilborn, E. D., J. H. Mermin, P. A. Mshar, J. L. Hadler, A. Voetsch, C. Wojtkunski, M. Swartz, R. Mshar, M. A. Lambert-Fair, J. A. Farrar, M. K. Glynn, and L. Slutsker.** 1999. A multistate outbreak of *Escherichia coli* O157:H7 infections associated with consumption of mesclun lettuce. Arch Intern Med **159**:1758-64.

53. **Hilborn, E. D., P. A. Mshar, T. R. Fiorentino, Z. F. Dembek, T. J. Barrett, R. T. Howard, and M. L. Cartter.** 2000. An outbreak of *Escherichia coli* O157:H7 infections and haemolytic uraemic syndrome associated with consumption of unpasteurized apple cider. Epidemiol Infect **124**:31-6.

54. **Hommais, F., E. Krin, J.-Y. Coppee, C. Lacroix, E. Yeramian, A. Danchin, and P. Bertin.** 2004. GadE (YhiE): a novel activator involved in the response to acid environment in *Escherichia coli.* Microbiology **150**:61-72.

55. **Hommais, F., E. Krin, C. Laurent-Winter, O. Soutourina, A. Malpertuy, J.-P. Le Caer, A. Danchin, and P. Bertin.** 2001. Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. Molecular Microbiology **40**:20-36.

56. **Ide, T., S. Michgehl, S. Knappstein, G. Heusipp, and M. A. Schmidt.** 2003. Differential modulation by Ca2+ of type III secretion of diffusely adhering enteropathogenic *Escherichia coli.* Infect Immun **71**:1725-32.

57. **Inward, C. D., A. J. Howie, M. M. Fitzpatrick, F. Rafaat, D. V. Milford, and C. M. Taylor.** 1997. Renal histopathology in fatal cases of diarrhoea-associated haemolytic uraemic syndrome. Pediatr Nephrol **11**:556-9.

58. **Issartel, J. P., V. Koronakis, and C. Hughes.** 1991. Activation of *Escherichia coli* prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. Nature **351:**759-61.

59. **Iyer, R., C. Williams, and C. Miller.** 2003. Arginine-agmatine antiporter in extreme acid resistance in *Escherichia coli*. J Bacteriol **185:**6556-61.

60. **Iyoda, S., and H. Watanabe.** 2004. Positive effects of multiple *pch* genes on expression of the locus of enterocyte effacement genes and adherence of enterohaemorrhagic *Escherichia coli* O157 : H7 to HEp-2 cells. Microbiology **150:**2357-2571.

61. **Iyoda, S., and H. Watanabe.** 2004. Positive effects of multiple *pch* genes on expression of the locus of enterocyte effacement genes and adherence of enterohaemorrhagic *Escherichia coli* O157 : H7 to HEp-2 cells. Microbiology **150:**2357-571.

62. **Jackson, S. G., R. B. Goodbrand, R. P. Johnson, V. G. Odorico, D. Alves, K. Rahn, J. B. Wilson, M. K. Welch, and R. Khakhria.** 1998. *Escherichia coli* O157:H7 diarrhoea associated with well water and infected cattle on an Ontario farm. Epidemiol Infect **120:**17-20.

63. **Jin, W., R. M. Riley, R. D. Wolfinger, K. P. White, G. Passador-Gurgel, and G. Gibson.** 2001. The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. Nat Genet **29:**389-395.

64. **Just, J. R., and M. A. Daeschel.** 2003. Antimicrobial Effects of Wine on *Escherichia coli* O157:H7 and *Salmonella typhimurium* in a Model Stomach System. Journal of Food Science **68:**285-290.

65. **Kailasan Vanaja, S., T. M. Bergholz, and T. S. Whittam.** 2009. Characterization of the *Escherichia coli* O157:H7 Sakai GadE regulon. J. Bacteriol. **191:**1868-1877.

66. **Kaper, J. B., J. P. Nataro, and H. L. Mobley.** 2004. Pathogenic *Escherichia coli*. Nat Rev Microbiol **2:**123-40.

67. **Karch, H., P. I. Tarr, and M. Bielaszewska.** 2005. Enterohaemorrhagic *Escherichia coli* in human medicine. Int J Med Microbiol **295:**405-18.

68. **Keene, W. E., J. M. McAnulty, F. C. Hoesly, L. P. Williams, Jr., K. Hedberg, G. L. Oxman, T. J. Barrett, M. A. Pfaller, and D. W. Fleming.** 1994. A swimming-associated outbreak of hemorrhagic colitis caused by *Escherichia coli* O157:H7 and *Shigella sonnei*. N Engl J Med **331:**579-84.

69. **Kenny, B., A. Abe, M. Stein, and B. B. Finlay.** 1997. Enteropathogenic *Escherichia coli* protein secretion is induced in response to conditions similar to those in the gastrointestinal tract. Infect Immun **65:**2606-12.

70. **Kenny, B., and B. B. Finlay.** 1995. Protein secretion by enteropathogenic *Escherichia coli* is essential for transducing signals to epithelial cells. Proc Natl Acad Sci U S A **92:**7991-5.

71. **Kern, R., A. Malki, J. Abdallah, J. Tagourti, and G. Richarme.** 2007. *Escherichia coli* HdeB Is an Acid Stress Chaperone. J. Bacteriol. **189:**603-610.

72. **Kim, J., J. Nietfeldt, and A. K. Benson.** 1999. Octamer-based genome scanning distinguishes a unique subpopulation of *Escherichia coli* O157:H7 strains in cattle. Proc Natl Acad Sci U S A **96:**13288-93.

73. **Kress, W., H. Mutschler, and E. Weber-Ban.** 2007. Assembly pathway of an AAA+ Protein: tracking ClpA and ClpAP complex formation in real time. Biochemistry **46:**6183-6193.

74. **Laegreid, W. W., R. O. Elder, and J. E. Keen.** 1999. Prevalence of *Escherichia coli* O157:H7 in range beef calves at weaning. Epidemiol Infect **123:**291-8.

75. **Large, T. M., S. T. Walk, and T. S. Whittam.** 2005. Variation in acid resistance among Shiga toxin-producing clones of pathogenic *Escherichia coli*. Appl. Environ. Microbiol. **71:**2493-2500.

76. **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:**277-88.

77. **Law, D.** 2000. Virulence factors of *Escherichia coli* O157 and other Shiga toxin-producing *E. coli*. J Appl Microbiol **88:**729-45.

78. **Law, D.** 2000. Virulence factors of *Escherichia coli* O157 and other Shiga toxin-producing *E. coli*. Journal of Applied Microbiology **88:**729-745.

79. **Lefebvre, B., M. S. Diarra, H. Moisan, and F. Malouin.** 2008. Detection of virulence-associated genes in *Escherichia coli* O157 and non-O157 isolates from beef cattle, humans, and chickens. J Food Prot **71:**1774-1784.

80. **Li, H., C. Wood, T. Getchell, M. Getchell, and A. Stromberg.** 2004. Analysis of oligonucleotide array experiments with repeated measures using mixed models. BMC Bioinformatics **5:**209.

81. **Lim, J. Y., H. Sheng, K. S. Seo, Y. H. Park, and C. J. Hovde.** 2007. Characterization of an *Escherichia coli* O157:H7 plasmid O157 deletion mutant and its survival and persistence in cattle. Appl Environ Microbiol **73:**2037-47.

82. **Lin, J., I. S. Lee, J. Frey, J. L. Slonczewski, and J. W. Foster.** 1995. Comparative analysis of extreme acid survival in *Salmonella typhimurium*, *Shigella flexneri*, and *Escherichia coli*. J Bacteriol **177:**4097-104.

83. **Lin, J., M. Smith, K. Chapin, H. Baik, G. Bennett, and J. Foster.** 1996. Mechanisms of acid resistance in enterohemorrhagic *Escherichia coli*. Appl. Environ. Microbiol. **62:**3094-3100.

84. **Ma, Z., S. Gong, H. Richard, D. L. Tucker, T. Conway, and J. W. Foster.** 2003. GadE (YhiE) activates glutamate decarboxylase-dependent acid resistance in Escherichia coli K-12. Molecular Microbiology **49:**1309-1320.

85. **Ma, Z., N. Masuda, and J. W. Foster.** 2004. Characterization of EvgAS-YdeO-GadE Branched Regulatory Circuit Governing Glutamate-Dependent Acid Resistance in *Escherichia coli*. J. Bacteriol. **186:**7378-7389.

86. **Ma, Z., H. Richard, and J. W. Foster.** 2003. pH-Dependent modulation of cyclic AMP levels and GadW-dependent repression of RpoS affect synthesis of the GadX regulator and *Escherichia coli* acid resistance. J Bacteriol **185:**6852-9.

87. **Ma, Z., H. Richard, D. L. Tucker, T. Conway, and J. W. Foster.** 2002. Collaborative Regulation of Escherichia coli Glutamate-Dependent Acid Resistance by Two AraC-Like Regulators, GadX and GadW (YhiW). J. Bacteriol. **184:**7001-7012.

88. **Manning, S. D., A. S. Motiwala, A. C. Springman, W. Qi, D. W. Lacher, L. M. Ouellette, J. M. Mladonicky, P. Somsel, J. T. Rudrik, S. E. Dietrich, W. Zhang, B. Swaminathan, D. Alland, and T. S. Whittam.** 2008. Variation in virulence among clades of *Escherichia coli* O157:H7 associated with disease outbreaks. Proc Natl Acad Sci U S A **105:**4868-73.

89. **Masuda, N., and G. M. Church.** 2002. *Escherichia coli* Gene Expression Responsive to Levels of the Response Regulator EvgA. J. Bacteriol. **184:**6225-6234.

90. **Masuda, N., and G. M. Church.** 2003. Regulatory network of acid resistance genes in *Escherichia coli.* Molecular Microbiology **48:**699-712.

91. **Mates, A. K., A. K. Sayed, and J. W. Foster.** 2007. Products of the *Escherichia coli* Acid Fitness Island Attenuate Metabolite Stress at Extremely Low pH and Mediate a Cell Density-Dependent Acid Resistance. J. Bacteriol. **189:**2759-2768.

92. **McConnell, E. L., A. W. Basit, and S. Murdan.** 2008. Measurements of rat and mouse gastrointestinal pH, fluid and lymphoid tissue, and implications for in-vivo experiments. Journal of Pharmacy and Pharmacology **60:**63-70.

93. **McDaniel, T., K. Jarvis, M. Donnenberg, and J. Kaper.** 1995. A Genetic Locus of Enterocyte Effacement Conserved Among Diverse Enterobacterial Pathogens. Proceedings of the National Academy of Sciences **92:**1664-1668.

94. **McDaniel, T. K., K. G. Jarvis, M. S. Donnenberg, and J. B. Kaper.** 1995. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. Proc Natl Acad Sci U S A **92:**1664-8.

95. **Mellies, J. L., A. M. S. Barron, and A. M. Carmona.** 2007. Enteropathogenic and Enterohemorrhagic Escherichia coli Virulence Gene Regulation. Infect. Immun. **75:**4199-4210.

96. **Michino, H., K. Araki, S. Minami, S. Takaya, N. Sakai, M. Miyazaki, A. Ono, and H. Yanagawa.** 1999. Massive outbreak of *Escherichia coli* O157:H7 infection in schoolchildren in Sakai City, Japan, associated with consumption of white radish sprouts. Am J Epidemiol **150:**787-96.

97. **Mundy, R., F. Girard, A. J. FitzGerald, and G. Frankel.** 2006. Comparison of colonization dynamics and pathology of mice infected with enteropathogenic *Escherichia coli*, enterohaemorrhagic *E. coli* and *Citrobacter rodentium.* FEMS Microbiol Lett **265:**126-32.

98. **Murphy, K., and K. Campellone.** 2003. Lambda Red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic *E. coli.* BMC Molecular Biology **4:**11.

99. **Nadler, C., Y. Shifrin, S. Nov, S. Kobi, and I. Rosenshine.** 2006. Characterization of Enteropathogenic *Escherichia coli* Mutants That Fail

To Disrupt Host Cell Spreading and Attachment to Substratum. Infect. Immun. **74:**839-849.

100. **Nakanishi, N., H. Abe, Y. Ogura, T. Hayashi, K. Tashiro, S. Kuhara, N. Sugimoto, and T. Tobe.** 2006. ppGpp with DksA controls gene expression in the locus of enterocyte effacement (LEE) pathogenicity island of enterohaemorrhagic *Escherichia coli* through activation of two virulence regulatory genes. Molecular Microbiology **61:**194-205.

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

102. **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:**1505-12.

103. **Oh, D. H., Y. Pan, E. Berry, M. Cooley, R. Mandrell, and F. Breidt, Jr.** 2009. *Escherichia coli* O157:H7 strains isolated from environmental sources differ significantly in acetic acid resistance compared with human outbreak strains. J Food Prot **72:**503-9.

104. **Oliveira, M. G., J. R. Brito, T. A. Gomes, B. E. Guth, M. A. Vieira, Z. V. Naves, T. M. Vaz, and K. Irino.** 2008. Diversity of virulence profiles of Shiga toxin-producing *Escherichia coli* serotypes in food-producing animals in Brazil. Int J Food Microbiol **127:**139-46.

105. **Opdyke, J. A., J.-G. Kang, and G. Storz.** 2004. GadY, a Small-RNA Regulator of Acid Response Genes in *Escherichia coli.* J. Bacteriol. **186:**6698-6705.

106. **Oporto, B., J. I. Esteban, G. Aduriz, R. A. Juste, and A. Hurtado.** 2008. *Escherichia coli* O157:H7 and non-O157 Shiga toxin-producing *E. coli* in healthy cattle, sheep and swine herds in Northern Spain. Zoonoses Public Health **55:**73-81.

107. **Osborne, M. J., N. Siddiqui, P. Iannuzzi, and K. Gehring.** 2004. The solution structure of ChaB, a putative membrane ion antiporter regulator from *Escherichia coli.* BMC Struct Biol **4:**9.

108. **Park, S., X. You, and J. A. Imlay.** 2005. Substantial DNA damage from submicromolar intracellular hydrogen peroxide detected in Hpx- mutants of *Escherichia coli.* Proc Natl Acad Sci U S A **102:**9317-9322.

109. **Paton, J. C., and A. W. Paton.** 1998. Pathogenesis and diagnosis of Shiga toxin-producing *Escherichia coli* infections. Clin Microbiol Rev **11**:450-79.

110. **Perna, N. T., G. Plunkett, 3rd, 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**:529-33.

111. **Persson, O., A. Valadi, T. Nystrom, and A. Farewell.** 2007. Metabolic control of the *Escherichia coli* universal stress protein response through fructose-6-phosphate. Mol Microbiol **65**:968-78.

112. **Peterson, W. L., P. A. Mackowiak, C. C. Barnett, M. Marling-Cason, and M. L. Haley.** 1989. The human gastric bactericidal barrier: mechanisms of action, relative antibacterial activity, and dietary influences. J Infect Dis **159**:979-83.

113. **Pfaffl, M. W.** 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucl. Acids Res. **29**:e45.

114. **Prevention, C. f. D. C. a.** 2007. *Escherichia coli* O157:H7 infection associated with drinking raw milk--Washington and Oregon, November-December 2005. MMWR Morb Mortal Wkly Rep **56**:165-7.

115. **Prevention, C. f. D. C. a.** 1997. *Escherichia coli* O157:H7 infections associated with eating a nationally distributed commercial brand of frozen ground beef patties and burgers--Colorado, 1997.

116. **Prevention, C. f. D. C. a.** 1995. *Escherichia coli* O157:H7 outbreak linked to commercially distributed dry-cured salami--Washington and California, 1994. MMWR Morb Mortal Wkly Rep **44**:157-60.

117. **Prevention, C. f. D. C. a.** 2006. Ongoing multistate outbreak of *Escherichia coli* serotype O157:H7 infections associated with consumption of fresh spinach--United States, September 2006. MMWR Morb Mortal Wkly Rep **55**:1045-6.

118. **Prevention, C. f. D. C. a.** 1993. Update: multistate outbreak of *Escherichia coli* O157:H7 infections from hamburgers--western United States, 1992-1993. MMWR Morb Mortal Wkly Rep **42**:258-63.

119. **Price, S. B., J. C. Wright, F. J. DeGraves, M.-P. Castanie-Cornet, and J. W. Foster.** 2004. Acid Resistance Systems Required for Survival of *Escherichia coli* O157:H7 in the Bovine Gastrointestinal Tract and in Apple Cider Are Different. Appl. Environ. Microbiol. **70:**4792-4799.

120. **Rangel, J. M., P. H. Sparling, C. Crowe, P. M. Griffin, and D. L. Swerdlow.** 2005. Epidemiology of *Escherichia coli* O157:H7 outbreaks, United States, 1982-2002. Emerg Infect Dis **11:**603-9.

121. **Reid, S. D., C. J. Herbelin, A. C. Bumbaugh, R. K. Selander, and T. S. Whittam.** 2000. Parallel evolution of virulence in pathogenic *Escherichia coli*. Nature **406:**64-7.

122. **Reitzer, L., and B. L. Schneider.** 2001. Metabolic context and possible physiological themes of sigma(54)-dependent genes in *Escherichia coli*. Microbiol Mol Biol Rev **65:**422-44.

123. **Richard, H., and J. W. Foster.** 2004. *Escherichia coli* Glutamate- and Arginine-Dependent Acid Resistance Systems Increase Internal pH and Reverse Transmembrane Potential. J. Bacteriol. **186:**6032-6041.

124. **Robinson, C. M., J. F. Sinclair, M. J. Smith, and A. D. O'Brien.** 2006. Shiga toxin of enterohemorrhagic *Escherichia coli* type O157:H7 promotes intestinal colonization. Proc Natl Acad Sci U S A **103:**9667-72.

125. **Rozen, S., and H. Skaletsky.** 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol **132:**365-86.

126. **Ruiz, C., L. M. McMurry, and S. B. Levy.** 2008. Role of the Multidrug Resistance Regulator MarA in Global Regulation of the *hdeAB* Acid Resistance Operon in *Escherichia coli*. J. Bacteriol. **190:**1290-1297.

127. **Russell, J. B., F. Diez-Gonzalez, and G. N. Jarvis.** 2000. Potential effect of cattle diets on the transmission of pathogenic *Escherichia coli* to humans. Microbes Infect **2:**45-53.

128. **Safdar, N., A. Said, R. E. Gangnon, and D. G. Maki.** 2002. Risk of hemolytic uremic syndrome after antibiotic treatment of *Escherichia coli* O157:H7 enteritis: a meta-analysis. Jama **288:**996-1001.

129. **Saitoh, T., S. Iyoda, S. Yamamoto, Y. Lu, K. Shimuta, M. Ohnishi, J. Terajima, and H. Watanabe.** 2008. Transcription of the *ehx* enterohemolysin gene is positively regulated by GrlA, a global regulator encoded within the locus of enterocyte effacement in enterohemorrhagic *Escherichia coli*. J Bacteriol **190:**4822-30.

130. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol **4:**406-25.

131. **Sayed, A. K., and J. W. Foster.** 2009. A 750 bp sensory integration region directs global control of the *Escherichia coli* GadE acid resistance regulator. Mol Microbiol **71:**1435-50.

132. **Sayed, A. K., C. Odom, and J. W. Foster.** 2007. The *Escherichia coli* AraC-family regulators GadX and GadW activate *gadE*, the central activator of glutamate-dependent acid resistance. Microbiology **153:**2584-2592.

133. **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.

134. **Schmittgen, T. D., and K. J. Livak.** 2008. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc **3:**1101-8.

135. **Scott, D. R., E. A. Marcus, D. L. Weeks, A. Lee, K. Melchers, and G. Sachs.** 2000. Expression of the *Helicobacter pylori urel* gene is required for acidic pH activation of cytoplasmic urease. Infect Immun **68:**470-7.

136. **Serra-Moreno, R., J. Jofre, and M. Muniesa.** 2007. Insertion site occupancy by stx2 bacteriophages depends on the locus availability of the host strain chromosome. J Bacteriol **189:**6645-54.

137. **Shao, C., Q. Zhang, W. Tang, W. Qu, Y. Zhou, Y. Sun, H. Yu, and J. Jia.** 2008. The changes of proteomes components of *Helicobacter pylori* in response to acid stress without urea. J Microbiol **46:**331-7.

138. **Shin, S., M.-P. Castanie-Cornet, J. W. Foster, J. A. Crawford, C. Brinkley, and J. B. Kaper.** 2001. An activator of glutamate decarboxylase genes regulates the expression of enteropathogenic Escherichia coli virulence genes through control of the plasmid-encoded regulator, Per. Molecular Microbiology **41:**1133-1150.

139. **Shin, S., M.-P. Castanie-Cornet, J. W. Foster, J. A. Crawford, C. Brinkley, and J. B. Kaper.** 2001. An activator of glutamate decarboxylase genes regulates the expression of enteropathogenic *Escherichia coli* virulence genes through control of the plasmid-encoded regulator, Per. Mol. Microbiol. **41:**1133-1150.

140. **Smith, D. K., T. Kassam, B. Singh, and J. F. Elliott.** 1992. *Escherichia coli* has two homologous glutamate decarboxylase genes that map to distinct loci. J. Bacteriol. **174:**5820-5826.

141. **Sperandio, V.** 2004. Striking a balance: inter-kingdom cell-to-cell signaling, friendship or war? Trends Immunol **25:**505-7.

142. **Sperandio, V., A. G. Torres, B. Jarvis, J. P. Nataro, and J. B. Kaper.** 2003. Bacteria-host communication: the language of hormones. Proc Natl Acad Sci U S A **100:**8951-6.

143. **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.

144. **Stewart, N., J. Feng, X. Liu, D. Chaudhuri, J. W. Foster, M. Drolet, and Y.-C. Tse-Dinh.** 2005. Loss of topoisomerase I function affects the RpoS-dependent and GAD systems of acid resistance in *Escherichia coli.* Microbiology **151:**2783-2791.

145. **Stim-Herndon, K. P., T. M. Flores, and G. N. Bennett.** 1996. Molecular characterization of *adiY*, a regulatory gene which affects expression of the biodegradative acid-induced arginine decarboxylase gene (*adiA*) of *Escherichia coli.* Microbiology **142 (Pt 5):**1311-20.

146. **Storey, J. D., and R. Tibshirani.** 2003. Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences **100:**9440-9445.

147. **Strachan, N. J., D. R. Fenlon, and I. D. Ogden.** 2001. Modelling the vector pathway and infection of humans in an environmental outbreak of *Escherichia coli* O157. FEMS Microbiol Lett **203:**69-73.

148. **Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P. Mesirov.** 2005. From the Cover: Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences **102:**15545-15550.

149. **Sung, L. M., M. P. Jackson, A. D. O'Brien, and R. K. Holmes.** 1990. Transcription of the Shiga-like toxin type II and Shiga-like toxin type II variant operons of *Escherichia coli.* J Bacteriol **172:**6386-95.

150. **Takada, A., G. Umitsuki, K. Nagai, and M. Wachi.** 2007. RNase E is required for induction of the glutamate-dependent acid resistance system in *Escherichia coli*. Biosci Biotechnol Biochem **71:**158-64.

151. **Tamura, K., J. Dudley, M. Nei, and S. Kumar.** 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol **24:**1596-9.

152. **Taneike, I., H. M. Zhang, N. Wakisaka-Saito, and T. Yamamoto.** 2002. Enterohemolysin operon of Shiga toxin-producing *Escherichia coli*: a virulence function of inflammatory cytokine production from human monocytes. FEBS Lett **524:**219-24.

153. **Tarr, P. I., C. A. Gordon, and W. L. Chandler.** 2005. Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. Lancet **365:**1073-86.

154. **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.

155. **Tatsuno, I., K. Nagano, K. Taguchi, L. Rong, H. Mori, and C. Sasakawa.** 2003. Increased adherence to Caco-2 cells caused by disruption of the *yhiE* and *yhiF* genes in enterohemorrhagic *Escherichia coli* O157:H7. Infect Immun **71:**2598-606.

156. **Tatsuno, I., K. Nagano, K. Taguchi, L. Rong, H. Mori, and C. Sasakawa.** 2003. Increased Adherence to Caco-2 Cells Caused by Disruption of the yhiE and yhiF Genes in Enterohemorrhagic Escherichia coli O157:H7. Infect. Immun. **71:**2598-2606.

157. **Tobe, T., H. Ando, H. Ishikawa, H. Abe, K. Tashiro, T. Hayashi, S. Kuhara, and N. Sugimoto.** 2005. Dual regulatory pathways integrating the RcsC-RcsD-RcsB signalling system control enterohaemorrhagic *Escherichia coli* pathogenicity. Molecular Microbiology **58:**320-333.

158. **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. Proceedings of the National Academy of Sciences **103:**14941-14946.

159. **Tramonti, A., M. De Canio, I. Delany, V. Scarlato, and D. De Biase.** 2006. Mechanisms of Transcription Activation Exerted by GadX and

GadW at the *gadA* and *gadBC* Gene Promoters of the Glutamate-Based Acid Resistance System in *Escherichia coli*. J. Bacteriol. **188**:8118-8127.

160. **Tsai, H. M., W. L. Chandler, R. Sarode, R. Hoffman, S. Jelacic, R. L. Habeeb, S. L. Watkins, C. S. Wong, G. D. Williams, and P. I. Tarr.** 2001. von Willebrand factor and von Willebrand factor-cleaving metalloprotease activity in *Escherichia coli* O157:H7-associated hemolytic uremic syndrome. Pediatr Res **49**:653-9.

161. **Tucker, D. L., N. Tucker, and T. Conway.** 2002. Gene Expression Profiling of the pH Response in *Escherichia coli*. J. Bacteriol. **184**:6551-6558.

162. **Tucker, D. L., N. Tucker, Z. Ma, J. W. Foster, R. L. Miranda, P. S. Cohen, and T. Conway.** 2003. Genes of the GadX-GadW Regulon in *Escherichia coli*. J. Bacteriol. **185**:3190-3201.

163. **Tuttle, J., T. Gomez, M. P. Doyle, J. G. Wells, T. Zhao, R. V. Tauxe, and P. M. Griffin.** 1999. Lessons from a large outbreak of *Escherichia coli* O157:H7 infections: insights into the infectious dose and method of widespread contamination of hamburger patties. Epidemiol Infect **122**:185-92.

164. **Umanski, T., I. Rosenshine, and D. Friedberg.** 2002. Thermoregulated expression of virulence genes in enteropathogenic *Escherichia coli*. Microbiology **148**:2735-2744.

165. **Van Winden, S. C., K. E. Muller, R. Kuiper, and J. P. Noordhuizen.** 2002. Studies on the pH value of abomasal contents in dairy cows during the first 3 weeks after calving. J Vet Med A Physiol Pathol Clin Med **49**:157-60.

166. **Wagner, W., M. Vogel, and W. Goebel.** 1983. Transport of hemolysin across the outer membrane of *Escherichia coli* requires two functions. J Bacteriol **154**:200-10.

167. **Waterman, S. R., and P. L. C. Small.** 2003. Transcriptional Expression of *Escherichia coli* Glutamate-Dependent Acid Resistance Genes *gadA* and *gadBC* in an *hns rpoS* Mutant. J. Bacteriol. **185**:4644-4647.

168. **Weber, A., S. A. Kogl, and K. Jung.** 2006. Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in *Escherichia coli*. J. Bacteriol. **188**:7165-7175.

169. **Weber, H., T. Polen, J. Heuveling, V. F. Wendisch, and R. Hengge.** 2005. Genome-Wide Analysis of the General Stress Response Network in

Escherichia coli: {sigma}S-Dependent Genes, Promoters, and Sigma Factor Selectivity. J. Bacteriol. **187:**1591-1603.

170. **Wick, L. M., W. Qi, D. W. Lacher, and T. S. Whittam.** 2005. Evolution of genomic content in the stepwise emergence of *Escherichia coli* O157:H7. J Bacteriol **187:**1783-91.

171. **Wick, L. M., W. Qi, D. W. Lacher, and T. S. Whittam.** 2005. Evolution of Genomic Content in the Stepwise Emergence of *Escherichia coli* O157:H7. J. Bacteriol. **187:**1783-1791.

172. **Wong, C. S., S. Jelacic, R. L. Habeeb, S. L. Watkins, and P. I. Tarr.** 2000. The risk of the hemolytic-uremic syndrome after antibiotic treatment of *Escherichia coli* O157:H7 infections. N Engl J Med **342:**1930-6.

173. **Yamanaka, K., W. Zheng, E. Crooke, Y. H. Wang, and M. Inouye.** 2001. CspD, a novel DNA replication inhibitor induced during the stationary phase in *Escherichia coli*. Mol Microbiol **39:**1572-84.

174. **Yuen, T., E. Wurmbach, R. L. Pfeffer, B. J. Ebersole, and S. C. Sealfon.** 2002. Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. Nucl. Acids Res. **30:**e48.

175. **Yuryev, A.** 2007. PCR primer design using statistical modeling. Methods Mol Biol **402:**93-104.

176. **Zwir, I., D. Shin, A. Kato, K. Nishino, T. Latifi, F. Solomon, J. M. Hare, H. Huang, and E. A. Groisman.** 2005. Dissecting the PhoP regulatory network of *Escherichia coli* and *Salmonella enterica*. Proc Natl Acad Sci U S A **102:**2862-7.