Use of signature-tagged mutagenesis to identify genes associated with colonization of sheep by *E. coli* O157:H7

Nancy A Cornick[#], Josh Pitzer^{*}, Amy F Helgersen[§], Melissa L. Madsen^a, Kathy T. Kurth[§], Qianjun Xiao^{**}, and F. Chris Minion

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

[#]Corresponding author

2130 Vet Med Bldg.
Iowa State University
Ames, IA 50010
PH 515-294-6499
FAX 515-294-8500
ncornick@iastate.edu

Running title – Colonization of ruminants by *E. coli* O157:H7

^{*}Present address. Josh Pitzer. Department of Microbiology and Immunology, East Carolina University, Greenville, NC 27834

[§]Present address. Amy Helgersen. Environmental Health and Safety, Iowa State Univ., Ames, IA 50011

^aPresent address. Melissa Madsen. CEVA Biomune, Lenexa, KS 66215

[§]Present address. Kathy T. Kurth. Wisconsin Veterinary Diagnostic Laboratory, University of Wisconsin-Madison, Madison, WI 53706

^{**}Present address. Qianjun Xiao. Neogen GeneSeek, Lincoln NE 68504

Abstract

Outbreaks of *Escherichia coli* O157:H7 in the United States due to contaminated foods are a public health issue and a continuing problem. The major reservoir for these organisms is the gastrointestinal tract of ruminants where they are a member of the resident microbiota. Several factors that contribute to the colonization of cattle have been

identified, but a systematic screen of genes that might contribute to the colonization and persistence phenotype in mature ruminants has not been reported. Using a sheep model of persistence, signature tagged mutagenesis (STM) was used to screen 1326 mutants for a persistence-negative phenotype of *E. coli* O157:H7. We identified 9 genes by STM that appeared to be required for colonization and/or survival in sheep. Three of the genes had functions associated with central metabolism (*thiK*, *ftrA* and *nrdB*), one was involved with LPS formation (*wbdP*), one encodes a non-LEE encoded effector protein (*nleB*) and one was a methyltransferase encoded on a prophage (Z2389). The remaining three genes did not have homology with any known genes. Six sheep given $\Delta wbdP$ and 2 sheep each were given mutants ($\Delta thiK$ (Z1745), $\Delta ftrA$ (Z2164) and Z2389). The $\Delta wbdP$ mutant was recovered from the feces of 4/6 sheep at 6 days pi with a mean number of $1.42 \log_{10}$ CFU/g feces compared to $4.6 \log_{10}$ CFU/g feces for the wild type strain. This difference was significant ($P < 0.001$) over the time course of the experiment (days 6 – 23). Both $\Delta thiK$ and $\Delta ftrA$ mutants were recovered from 1 of 2 sheep at 9 days PI by enrichment procedures (< 50 CFU/g feces) whereas mutant Z2389 was not recovered from either animal past 2 days pi. The roles of all of these gene products require further study to determine how the persistence phenotype of a given strain of *E. coli* O157:H7 interacts with host factors.

Introduction

Shiga toxin-producing *Escherichia coli* (STEC) are an important cause of food-borne illness worldwide, and ruminants are considered the major reservoir of these organisms. STEC are not pathogenic in adult cattle or sheep and can be shed from the gastrointestinal tract for several months (Cray and Moon, 1995). It is estimated that 90% of cattle herds are colonized by one or more serotypes of STEC (Wells et al., 2009). In the US, *E. coli* O157:H7 is the serotype most often associated with human disease. Most cattle shed $< 10^2$ CFU/g of *E. coli* O157:H7 in their feces, while super-shedders are defined as those animals that shed $> 10^4$ CFU/g (Matthews et al., 2006). Since the infectious dose of STEC for humans is typically low ($< 10^2$ bacteria have been reported), the presence of super-shedders pose an increased risk to human health following incidental contact with contaminated food or water sources (Matthews et al., 2006). In addition, some studies have suggested that super-shedders may be colonized for a longer period of time compared to cohorts that are shedding fewer organisms or those that are transiently colonized (Naylor et al., 2003; Rice et al., 2003). Alternatively, low-level persistent colonization, interspersed with clonal population blooms, may account for the

‘intermittent’ shedding that is frequently detected in ruminants (Schlager et al., 2002; Whittam, 1989). Current limitations on culture detection, even with immunomagnetic concentration, may not be sensitive enough to detect extremely low-level populations of *E. coli* O157:H7 that could continue to serve as a seed for clonal expansion.

Only a few genes have been shown to be critical for the colonization of ruminants by *E. coli* O157:H7, the prototype STEC. These genes are *eae*, which encodes the outer membrane protein intimin (important for binding to epithelial cells); *tir* which encodes the translocated intimin receptor; and *espA* which encodes the filament component of the type III secretion system (Cornick et al., 2002; Dziva et al., 2004). All three of these genes are found on the large pathogenicity island, the locus of enterocyte effacement (LEE). Additional adhesive factors that appear to modulate intestinal colonization of ruminants include the H7 flagella (Dobbin et al., 2006), *lpf* fimbriae (Jordan et al., 2004), *stcE*, an esterase inhibitor (Eckert et al., 2011), *espK*, an effector protein (Vlisidou et al., 2006) and *espP*, a serine protease (Dziva et al., 2007). It is likely that initial binding to host enterocytes occurs by a mechanism other than the intimin-mediated binding encoded by the LEE genes. The recto anal junction is an important site of colonization for *E. coli* O157:H7 in cattle, but there is less evidence in sheep that this site is more readily colonized than the remainder of the large intestine (Cornick, personal observation). Some non-O157 STEC that do not contain a LEE pathogenicity island also colonize cattle (Orden et al., 2005) and cause serious disease in humans, suggesting that these strains must have alternate adherence mechanisms as well (Brooks et al., 2005).

Signature tagged mutagenesis (STM) is the application of individual sequence-tagged transposons to identify genes essential for growth *in vivo*. The power of STM is exemplified by its ability to characterize multiple mutants simultaneously and to identify the loss of specific DNA sequence tags (molecular bar codes) during *in vivo* passage (Hensel et al., 1995). The sequence tags within the mutant ‘input’ population are compared to those mutants recovered from the host, and those that represent mutations in essential *in vivo* genes fail to survive or are undetectable. Mutants missing from the recovered pool can be classified into several different groups including those that fail to colonize and are immediately eliminated from the host, and those that colonize, but fail to survive immune defense or environmental insults. These latter mutants persist for a short period of time, but are lost from the output pool much sooner than the wild type parent. Prior to the development of STM, this phenotype was difficult to study because of the need to examine

individual mutants in multiple animals. STM has been used previously to investigate STEC colonization in calves. However, these experiments were conducted in pre-ruminant calves that were evaluated only five days post-inoculation (Dziva et al., 2004; Eckert et al., 2011; Van Diemen et al., 2005). We were interested in using this approach to identify genes *in vivo* that would be associated with longer-term, more persistent colonization of mature ruminants, a model that more closely mimics the *E. coli* O157:H7 reservoir.

Materials and Methods.

Bacterial strains and growth conditions. A spontaneous nalidixic acid mutant of *E. coli* O157:H7 strain 86-24 (Nal^R), and *E. coli* K-12 strains *E. coli* CC118λpir and BW19795λpir were grown overnight at 37 °C in Luria Broth (LB) or on MacConkey agar. Antibiotics were added as appropriate at the following concentrations except where indicated: ampicillin (Amp, 100 µg/ml); kanamycin (Kan, 50 µg/ml); nalidixic acid (Nal, 20 µg/ml); and streptomycin (Str, 50 µg/ml). SOC medium (20 g tryptone, 5 g yeast extract, 2.03 g MgCl₂•6H₂O, 2.46 g MgSO₄•7H₂O, 3.6 g glucose per liter, pH 7.5) was used for freezing recovered mutants and GN Hajna broth (Difco) was used for enrichment cultures.

Construction of *E. coli* O157:H7 STM library. The mini-Tn5Km2 STM plasmid library was obtained from Brenden Wren (Karlyshev et al., 2001) as a set of 96 plasmids containing individual miniTn5Km2 transposons with double unique sequence tags on a Pir-dependent ampicillin-resistant plasmid. The plasmids were maintained individually in *E. coli* CC118λpir by selecting for ampicillin resistance, and then transformed into the conjugal donor strain BW19795λpirF⁺. This generated a library of 96 transformants, each with an individual sequence tag.

Conjugations were performed with the 96 BW19795λpirF⁺ transformants and *E. coli* O157 86-24 Nal^R as described using a membrane filter method (Naylor et al., 2003). Exconjugants were identified by plating conjugation mixes on Nal Kan LB plates. Each plate yielded 300-400 colonies of *E. coli* O157 86-24 STM mutants. One hundred ninety-two individual exconjugants from each mating mixture, each representing a single unique transposon insertion, were picked to microtiter plates and stored at -70 °C. Ampicillin resistant mutants, indicating integration of the plasmid, were identified by replica plating onto LB ampicillin media. These were subsequently discarded from further study. After the 96 conjugation experiments were completed, the individual conjugant groups were assigned to a location in a microtiter

plate and re-arrayed so that each member of a plate contained a unique, separate tag. Each plate then became an inoculum pool of 95 individual tagged transposon mutants. Well A7 served as a negative control in each pool.

***In vivo* experiments.** Young adult sheep were obtained from a local supplier, housed in ABSL-2 conditions and adapted to an alfalfa/hay mixture plus 2.2 kg of feed concentrate/animal for one week prior to inoculation. All of the animal procedures were approved by the ISU Institutional Committee for Animal Care and Use. For the initial round of screening, a pool of 95 or 47/48 signature tagged mutants (approximately 10^9 CFU per mutant) was prepared by replica plating each pool into a deep well microtiter plate containing 1.25 ml LB medium/well and incubating overnight at 37 °C with shaking. The mutants were combined, centrifuged at $6,500 \times g$ at 4 °C for 10 min. The pellet was re-suspended in Trypticase soy broth (TSB), one ml of the original inoculum pool was refrozen to be used as the input pool. The concentration of the inoculum was calculated by direct plating.

Two sheep were inoculated with 10^9 CFU of each inoculum pool of 95 or 47/48 STM mutants. A total of 36 sheep were used for the first round of screen and 10 sheep were used in the second round of screening. Sheep were inoculated by adding the pool of bacterial mutants to a small portion of feed concentrate. After the inoculum was consumed, animals were fed their remaining concentrate and hay. Fecal samples were collected from sheep on days 20, 22, 24 (3 weeks) post inoculation. Samples were collected after the rooms had been cleaned by penning animals individually for several hours and collecting fresh fecal pellets from the floor. Samples were serial diluted, inoculated onto MacConkey plates with and without antibiotics (Nal-Kan) in triplicate and incubated overnight at 37 °C. Enrichment broth containing Nal (10 µg/ml) and Kan was inoculated with 10 g of feces, incubated overnight and plated onto selective medium. If the 10^{-1} dilution of Nal-Kan MacConkey plate contained 10,000 colonies, bacterial DNA was collected directly from this plate. Otherwise 200 µl of enrichment broth was inoculated onto two plates of selective medium and incubated at 37 °C overnight. Colonies were confirmed by O157 latex agglutination reaction. Colonies were scraped into a 1.5 ml tube containing 10% glycerol in SOC broth. Samples were stored at -80 °C. The sheep were euthanized at the end of the experiment.

For the second round of screening, the clones that were missing from at least one of the sheep at 3 weeks post inoculation (pi) were re-arrayed into 5 new inoculum pools each containing 6-9 individual mutants (Fig 1). Two sheep

were inoculated with each pool as described above and fecal samples were cultured at 2 (days 13, 15, 17) and 4 weeks (days 27, 29, 31) pi. Colonies were scraped into a 1.5 ml tube containing 10% glycerol in SOC broth and stored at -80 °C.

DNA extraction and PCR. Samples were thawed on ice and pooled by week for each animal. Cells were pelleted using a microcentrifuge and resuspended in lysis buffer containing 10 mM Tris buffer, 10 mM EDTA, lysozyme (1 mg/ml) and 0.02% Triton X-100. Cells were incubated for 5 min at 37 °C followed by 10 min at 95 °C. Cellular debris was removed by centrifugation and the supernatant was used in PCR reactions. For detection of individual sequence tags, PCR was performed using a 25 µl reaction volume containing 2.5 µl 10X Thermo Pol Buffer (New England BioLabs) with 2 mM MgCl₂, 1 µl dNTP mix containing 1 mM each base, 1 µl primer AK84 (5'-CAC TGG CAG AGC ATT ACG CTG) at a concentration of 10 µM, 1 µl of sequence tag-specific primer (Table S1) at a concentration of 10 µM, 0.5 µl of Taq, 0.1 µl DNA template, 18.9 µl H₂O. The reaction was performed under the following conditions: 94 °C for 2 min then 30 cycles of 94 °C for 30 sec, 66 °C for 30 sec, 72 °C for 15 sec followed by an extension reaction of 72 °C for 2 min. Ten µl of PCR products were analyzed on 1.5% agarose gels and visualized with ethidium bromide. In order for a tag to be considered missing from the input or output pool, both tag1 and tag2 had to be missing for that particular transposon. Output pools were compared to input pools and colonization-negative mutants were defined as those that were present in the input pool, but absent from the output pool.

Sequencing of transposon mutants. To determine the transposon location, a single primer PCR was performed and the product was sequenced at the Iowa State University DNA facility using dideoxy sequencing chemistry. To perform the PCR, a 50 µl reaction volume containing 5 µl 10X Thermo Pol Buffer (New England BioLabs), 2 µl dNTP mix containing 1 mM each base, 0.5 µl Taq, 4 µl P6U (5'-CGAGCTCGAATTCGGCCTAG) or P7U (5'-CTGCAGGCATGCAAGCTTCG) primer at 10 µM, 1 µl template, 37.5 µl H₂O was run under the following conditions: 94 °C for 1 min and then 20 cycles of 94 °C for 30 sec; 50 °C for 30 sec, 72 °C for 3 min; followed by 30 cycles of 94 °C for 30 sec, 30 °C for 30 sec, 72 °C for 2 min; and then 30 cycles of 94 °C for 30 sec, 50 °C for 30 sec, 72 °C for 2 min followed by a 7 min extension at 72 °C as described (Karlyshev et al., 2000). The reaction products were analyzed on 1.5% agarose gels, and the products were purified to remove excess primers, salts, and dNTP's. Sequence reactions were performed with the PCR products, P6U reactions

with primer P6M (5' GCCAGATCTGATCAAGAGAC) and P7U reactions with primer P7M (5' GCCGAACTTGTGTATAAGAGTC).

Deletion mutants. Deletion mutants were constructed in *E. coli* O157:H7 86-24 in putative persistence-associated genes by homologous recombination (Datsenko and Wanner, 2000; Dziva et al., 2007). Briefly, a linear PCR product containing the Kan resistance cassette plus 30-50 bp of the DNA sequence immediately 3' and 5' of each gene of interest was amplified and then electroporated into *E. coli* O157:H7 strain 86-24 containing the temperature-sensitive λ red helper plasmid pRedET (Gene Bridges, Dresden, Germany). Table S2 lists the primers used in each reaction to generate the recombinering fragment. Transformants were recovered by incubation at 42 °C on TSA or LB agar plates containing 50 μ g/ml Kan. The recombinering fragment was designed so that the insertion of the antibiotic resistance cassette in the wild-type strain was in-frame. This was confirmed by PCR sequencing to ensure that only the selected gene was knocked out and no other downstream genes.

In vitro competitive growth assays. Cultures of individual strains were grown overnight in LB broth, diluted to $\sim 1 \times 10^5$ CFU/ml then combined together in a 1:1 ratio in fresh LB broth and incubated with shaking (200 rpm) for 6 h. Aliquots were removed at 0, 2, 4 and 6 h after mixing, serially diluted and plated onto selective media (MacConkey agar plates containing Strep 100 μ g/ml and Nal 20 μ g/ml for recovery of the wild type strain or MacConkey containing Kan 30 μ g/ml and Nal 20 μ g/ml to recover mutants. The specific growth rate was calculated using the formula $k = (t_2 - t_1) / 2.303$ (Miller, 2011). A t-test was used to compare the specific growth rate of each mutant grown alone and in the presence of the parent strain.

Adherence to Hep-2 cells. Quantitative assays to compare the adherence of the mutants with that of the wild type strain were performed as described previously (Ojea, 2012; Puttamreddy et al., 2010). Briefly, monolayers of Hep-2 cells in 24 well tissue culture plates were inoculated with approximately 4×10^6 CFU of each bacterial strain (mutant or wild type) and incubated for 6 h at 37 °C. Following incubation, the plates were washed 3 times with PBS to remove non-adherent bacteria. A 0.1% solution of Triton-X 100 was added (0.5 ml) to detach and lyse the remaining cells. Adherent bacteria were serially diluted and plated onto MacConkey agar with antibiotics and enumerated. Adherence was

calculated using the formula: number bacteria adhered / number bacteria added to the well x 100. A $\Delta ehxD$ mutant was used as a negative control (Puttamreddy et al., 2010). A paired t-test was used to compare differences between strains.

***In vivo* colonization by specific mutants.** Equal numbers ($\sim 10^{10}$ CFU) of the mutant and the isogenic wild type strains were orally inoculated into sheep by adding the strains to a small amount of feed as described above. Six sheep given $\Delta wbdP$ and 2 sheep each were given mutants $\Delta thiK$ (Z1745), $\Delta ftrA$ (Z2164) and Z2389. Fecal samples were collected 2 times per week for 4 weeks and cultured for both the mutant and wild type strains. Briefly, fecal samples were serially diluted (10-fold) and directly plated onto selective media with antibiotics. In addition, 10 g fecal samples were enriched by overnight incubation in 100 ml of GN broth (no antibiotics) then concentrated using immunomagnetic beads (Dynabeads) and plated onto selective agar. Presumptive *E. coli* O157:H7 colonies were confirmed using latex agglutination. Numbers of mutants were compared to the parent strain using repeated measures ANOVA modeling for overall differences and a signed rank test for comparisons on specific days.

Results

STM library construction. Conjugation and rearraying resulted in 100 microtiter plates containing individual transposon mutants. While each well contained a known sequence tag, i.e., the A1 well contained the A1 set of sequence tags, the exconjugate in each well contained a random transposon insertion from the original conjugation experiment. Sets of these 100 pools of mutants were randomly chosen for use in subsequent sheep experiments.

***In vivo* screening of STM library.** Ten pools of 95 mutants each and 8 pools of 47 or 48 mutants each were inoculated into two sheep each (Fig 1). In preliminary studies, the larger pools resulted in many missing members from each pool. Reducing the number of mutants from 95 to 48 resulted in more consistent detection of missing mutants from both animals (data not shown). From the 1,326 clones inoculated into sheep, 51 mutants (38 were unique) were not recovered from at least one of the animals at 3 weeks pi. This time period was selected for evaluation since it was well past the initial wash-out phase of the challenge and any recovered mutants would be expected to have established colonization in the gastrointestinal tract. The mean shedding of *E. coli* O157:H7 was 1.28×10^3 CFU/g feces at this time point. The 38 mutants that were not recovered on the initial screening were inoculated into 2 additional sheep (5

different pools, each with 6-9 mutants), and after this second round of screening, 9 mutants were missing from 3 of 4 samples collected (2 wk samples and 4 wk samples from each animal were pooled by week x 2 animals). The genes containing the transposon were identified by DNA sequencing using single strand PCR and these 9 genes are listed in Table 1. Three of the genes had functions associated with central metabolism, one was involved with LPS formation, one encodes a non-LEE encoded effector protein and one was encoded on a prophage. The remaining three genes did not have homology with any known genes.

Site directed mutants. Deletion mutations were constructed in 7 of the 9 genes identified by STM by recombineering as shown on Table 1. At this time, we have not successfully created deletion mutations in Z00065 or *nrdB* by recombineering despite several attempts.

Competitive growth curves. Specific growth rates of each deletion mutant grown alone and in competition with the parent strain are listed on Table 2. When the specific growth rates of the mutants were compared to that of the parent strain both Δbax and $\Delta wbdP$ grew at a significantly ($P \leq 0.05$) slower rate than the parent strain. The specific growth rates of $\Delta ftrA$ and $\Delta yihF$ were significantly less ($P \leq 0.05$) when these strains had to compete with the parent strain. In contrast, the growth rate of the Δbax mutant was increased when it was grown with the parent strain compared to its' growth alone. The growth rates of the other mutants were not impacted by the presence of the parent stain.

Adhesion assays. All of the mutants except Z2389 adhered to Hep-2 cells in equal or greater numbers than the wild type strain (Fig. 2). There were no significant differences between any of the mutant strains in their ability to bind to Hep-2 cells.

In vivo colonization. Four of the mutants (*thiK*, *ftrA*, Z2389, and *wbdP*) were randomly selected to test *in vivo* using our sheep model of colonization. The $\Delta wbdP$ mutant was recovered from the feces of 6/6 sheep at 2 days pi and from 4/6 sheep at 6 days pi (Fig 3). The mean fecal shedding at 6 days pi was $1.42 \log_{10}$ CFU/g feces for the $\Delta wbdP$ mutant compared to $4.6 \log_{10}$ CFU/g feces for the parent strain. This difference was significant ($P < 0.001$) over the time course of the experiment (days 6 – 23). The $\Delta wbdP$ mutant was not recovered after 6 days pi whereas the parent strain was recovered from 5/6 sheep at 27 days pi (<50 CFU/g feces). Three additional mutants, $\Delta thiK$ (Z1745), $\Delta ftrA$ (Z2164) and

Z2389, were inoculated into 2 sheep each. Both $\Delta thiK$ and $\Delta ftrA$ mutants were recovered from 1 of 2 sheep at 9 days PI by enrichment procedures (<50 CFU/g feces) (Fig. 4). Mutant Z2389 was not recovered from either animal past 2 days pi whereas the parent wild type strain was recovered from the feces of all 6 sheep (<50 CFU/g feces) at 27 days pi. The wild type strain was recovered from the tissues (cecum, spiral colon, descending colon) of 3/6 sheep at necropsy (30 days pi) whereas none of the mutants were recovered from any tissues.

Discussion

Using a STM screening approach, we identified mutations in nine genes that are required for colonization and/or survival of *E. coli* O157:H7 in the gastrointestinal tract of sheep. We randomly selected four of these mutants to test *in vivo* by inoculating them back into animals and following fecal shedding. The *wbdP* gene encodes a glycosyl transferase that is involved in modifying the surface LPS of the bacterium and has also been identified previously by Dziva *et al.* as being required for intestinal colonization in calves by *E. coli* O157:H7 (Dziva *et al.*, 2004). It is well known that cattle, and presumably sheep, make antibodies against the O157 LPS following experimental inoculation (Johnson *et al.*, 1996). However, it is not clear how the LPS is involved in host colonization. Results from *in vitro* binding studies using LPS mutants have been inconsistent, although some of the variation could be due to the use of different cell lines by various investigators (Bilge *et al.*, 1996; Cockerill *et al.*, 1996). In our hands, the Δwbd mutant adhered in significantly greater numbers to Hep-2 cells compared to the wild type parent, however this did not translate into colonization *in vivo* as measured by recovery of the strain from inoculated animals. We chose to use Hep-2 cells since they have been used to define binding patterns of different pathogenic *E. coli* strains. *E. coli* O157:H7 *gal* mutants, which do not make an O antigen, are defective in the ability to colonize the intestine of infant rabbits (Ho and Waldor, 2007). Likewise, deep rough mutants of a probiotic *E. coli* (Nissle strain) and those from an O18:K1:H7/9 strain associated with neonatal meningitis, were severely impaired with regards to *in vivo* colonization of mice (Nagy *et al.*, 2005). Further studies with these latter LPS mutants indicated that they adhered to colonic cells *in vitro*, but had an increased sensitivity to bile salts compared to the wild type parent strains. It may be that the decreased colonization of the $\Delta wbdP$ mutant in sheep is due to limited survival *in vivo*, rather than to a direct colonization defect.

It was not particularly surprising that mutants created in genes required for basic metabolic functions such as *thiK* and *ftrA* were not able to compete *in vivo* with the wild type strain. The *thiK* gene is required for thiamine metabolism and catalyzes the phosphorylation of thiamine to thiamine phosphate. The transcriptional regulator, *ftrA* is upregulated under O₂ limiting conditions and regulates genes involved in carbon and amino acid metabolism. The defects in these mutants may have more to do with decreased survival *in vivo* rather than a specific impairment to intestinal colonization. Indeed, the Δftr mutant grew significantly slower in a competitive setting *in vitro* (Table 2).

Z2389 is located within a prophage and encodes a DNA cytosine methyltransferase (Chiou et al., 2010). In general, methyltransferases are involved in restriction modification systems and serve to protect bacteria from insertion of exogenous DNA. Much less is known regarding the biological roles of orphan DNA methyltransferases, such as Z2389, and it is not apparent as to why the loss of this gene reduced colonization in sheep. However a study in *Mycobacterium tuberculosis* suggested that a DNA methyltransferase provided increased fitness *in vivo* in specific host environments (Shell et al., 2012). It is possible that this also occurs in *E. coli* O157:H7. Our results suggest that genes encoding enzymes and proteins of central metabolic pathways and basic bacterial cell maintenance may be more directly involved in aspects of colonization or survival *in vivo* than has been previously appreciated. Indeed within the last several years there has been renewed interest in the links between bacterial metabolism *in vivo* and virulence (Passalacqua et al., 2016)). We failed to identify any genes from the LEE pathogenicity island as putative persistence genes using our STM screening approach. This was somewhat surprising since we and others have shown that intimin, *tir* and *espA* are required for colonization of ruminants by *E. coli* O157:H7 (Cornick et al., 2002; Dziva et al., 2007; Dziva et al., 2004). Of the 10,000 recombinants created, this study examined 1,326 exconjugants. Screening greater numbers of exconjugants likely would have identified additional candidate genes involved in colonization and persistence.

Our results also confirm that *in vivo* screening methods, such as STM, are a critical component of identifying genes required for *in vivo* survival and metabolic potential of resident microbiota. The four hypothetical genes identified in this study are unique (Table 1) and have not been previously associated with persistence of *E. coli* O157:H7 in ruminants. Z4328 (*nleB*) is a non-LEE-encoded effector protein secreted by the type III secretion system of the mouse pathogen *Citrobacter rodentium* and the human EPEC strain 2348/69 (Deng et al., 2004; Kelly et al., 2006). This protein is

involved in colonization of mice and here we show that it is important for persistence of *E. coli* O157:H7 in ruminants. Although some studies have indicated that NleB may not be secreted under *in vitro* conditions (Roe et al., 2007), our data suggest that it is active in the ruminant intestinal tract. Other non-LEE encoded proteins have been shown to be required for adherence to stratified squamous epithelium of the bovine rectum (Kudva et al., 2012), so our finding is not without precedent. Additionally, in an independent study, using *E. coli* O157:H7 strain 933, we identified Z4328 (*nleB*), as well as, *wbdP* as being required for *in vitro* biofilm formation (Puttamreddy et al., 2010). We are currently investigating the relationship between biofilm formation and intestinal colonization. Of the remaining hypothetical genes we identified, Z5393 (*yihF*) is a putative GTP-binding protein with unknown function and Z4995 (*bax*) is a putative ATP-binding protein with unknown function. The final hypothetical gene, Z0065, is also without known function. Further study of the role of all of these gene products in vivo persistence/colonization will be necessary to fully understand the complexities of this phenotype.

Acknowledgements

We thank Chong Wang for the statistical analysis and gratefully acknowledge the assistance of Jan Seibel, Mary Jane Long, Chelsea Wetjen, Erin Brown, and Carissa Ralph for laboratory work and animal handling. This work was supported in part by grants from the USDA NRIGCP 2002-35201-11611 to FCM, ISU Colleges of Veterinary Medicine and Agriculture /Life Sciences to NC.

Literature cited

- Bilge, S.S., Vary, J.C.J., Dowell, S.F., Tarr, P.I., 1996. Role of the *Escherichia coli* O157:H7 O side chain in adherence and analysis of an *rfb* locus. Infect. Imm. 64, 4795-4801.
- Brooks, J.T., Sowers, E.G., Wells, J.G., Greene, K.D., Griffin, P.M., Hoekstra, R.M., Strockbine, N.A., 2005. Non-O157 shiga-toxin-producing *Escherichia coli* infections in the United States, 1983-2002. J. Infect. Dis. 192, 1422-1429.
- Chiou, C.S., Li, H.Y., Tung, S.K., Chen, C.Y., Teng, C.H., Shu, J.C., Tseng, J.T., Hsu, C.Y., Chen, C.C., 2010. Identification of prophage gene z2389 in *Escherichia coli* EDL933 encoding a DNA cytosine methyltransferase for full protection of NotI sites. J. Med. Microbiol. 300, 296-303.

- Cockerill, F.I., Beebakhee, G., Soni, R., Sherman, P.M., 1996. Polysaccharide side chains are not required for attaching and effacing adhesion of *Escherichia coli* O157:H7. *Infect. Imm.* 64, 3196-3200.
- Cornick, N.A., Booher, S., Moon, H.W., 2002. Intimin facilitates colonization by *Escherichia coli* O157:H7 in ruminants. *Infect. Imm.* 70, 2704-2707.
- Cray, W.C.J., Moon, H.W., 1995. Experimental infection of calves and adult cattle with *Escherichia coli* O157:H7. *Appl. Environ. Microbiol.* 61, 1586-1590.
- Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS* 97, 6640-6645.
- Deng, W., Puente, J.L., Gruenheid, S., Li, Y., Vallance, B.A., Vazquez, A., Barba, J., Ibarra, J.A., O'Donnell, P., Metalnikov, P., Ashman, K., Lee, S., Goode, D., Pawson, T., Finlay, B., 2004. Dissecting virulence: systematic and functional analyses of a pathogenicity island. *PNAS* 101, 3597-3602.
- Dobbin, H.S., Hovde, C.J., Williams, C.J., Minnich, S.A., 2006. The *Escherichia coli* O157 flagellar regulatory gene *flhC* and not the flagellin gene *fliC* impacts colonization of cattle. *Infect. Imm.* 74, 2894-2905.
- Dziva, F., Mahajan, A., Cameron, P., Currie, C., McKendrick, C., Wallis, T.S., Smith, D.G.E., Stevens, M.P., 2007. EspP, a type V-secreted serine protease of enterohemorrhagic *Escherichia coli* O157:H7, influences intestinal colonization of calves and adherence to bovine primary intestinal epithelial cells. *FEMS Microbiol. Lett.* 271, 258-264.
- Dziva, F., Van Diemen, P.M., Stevens, M.P., Smith, A.J., Wallis, T.S., 2004. Identification of *Escherichia coli* O157:H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis. *Microbiol.* 150, 3631-3645.
- Eckert, S.E., Dziva, F., Chaudhuri, R.R., Langridge, G.C., Turner, D.J., Pickard, D.J., Maskell, D.J., Thomson, N.R., Stevens, M.P., 2011. Retrospective application of transposon-directed insertion site sequencing to a library of signature-tagged mini-Tn5Km2 mutants of *Escherichia coli* O157:H7 screened in cattle. *J. Bacteriol.* 193, 1771-1776.
- Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E., Holden, D.W., 1995. Simultaneous identification of bacterial virulence genes by negative selection. *Science* 269, 400-403.

- Ho, T.D., Waldor, M.K., 2007. Enterohemorrhagic *Escherichia coli* O157:H7 *gal* mutants are sensitive to bacteriophage P1 and defective in intestinal colonization. *Infect. Imm.* 75, 1661-1666.
- Johnson, R.P., Cray, W.C.J., Johnson, S.T., 1996. Serum antibody responses of cattle following experimental infection with *Escherichia coli* O157:H7. *Infect. Imm.* 64, 1879-1883.
- Jordan, D.M., Cornick, N.A., Torres, A.G., Dean-Nystrom, E.A., Kaper, J.B., Moon, H.W., 2004. Long polar fimbriae contribute to colonization by *Escherichia coli* O157:H7 in vivo. *Infect. Imm.* 72, 6168-6171.
- Karlyshev, A.V., Oyston, P.C.F., Clark, C.G., Titball, R.W., Winzeler, E.A., Wren, B.W., 2001. Application of high-density array-based signature-tagged mutagenesis to discover novel *Yersinia* virulence-associated genes. *Infect. Imm.* 69, 7810-7819.
- Karlyshev, A.V., Pallen, M.J., Wren, B.W., 2000. Single-primer PCR procedure for rapid identification of transposon insertion sites. *Biotechniques* 28, 1080-1082.
- Kelly, M., Hart, E., Mundy, R., Marches, O., Wiles, S., Badea, L., Luck, S., Tauschek, M., Frankel, G., Robbins-Browne, R.M., Hartland, E.L., 2006. Essential role of the type III secretion system effector NleB in colonization of mice by *Citrobacter rodentium*. *Infect. Imm.* 74, 2328-2337.
- Kudva, I., Griffin, R.W., Krastins, B., Sarracino, D.A., Calderwood, S.B., John, M., 2012. Proteins other than the locus of enterocyte effacement-encoded proteins contribute to *Escherichia coli* O157:H7 adherence to bovine rectoanal junction stratified squamous epithelial cells. *BMC Microbiol.* 12, 103.
- Matthews, L., Low, C., Gally, D.L., Pearce, M.C., Mellor, D.J., Heesterbeek, J.A.P., Chase-Topping, M., Naylor, S.W., Shaw, D.J., Reid, S.D., Gunn, G.J., Woolhouse, M.E.J., 2006. Heterogeneous shedding of *Escherichia coli* O157 in cattle and its implications for control. *Proceeding of the National Academy of Sciences* 103, 547-552.
- Miller, S. 2011. General bacteriology protocols. <http://miller-lab.net/MillerLab/protocols/general-bacteriology/calculating-growth-rate/>
- Nagy, G., Dobrindt, U., Grozdanov, L., Hacker, J., Emody, L., 2005. Transcriptional regulation through RfaH contributes to intestinal colonization by *Escherichia coli*. *FEMS Microbiol. Lett.* 244, 173-180.

- Naylor, S.W., Low, J.C., Besser, T.E., Mahajan, A., Gunn, G.J., Pearce, M.C., McKendrick, I.J., Smith, D.G.E., Gally, D.L., 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. Imm.* 71, 1505-1512.
- Ojea, M.M.S., 2012. Contribution of biofilm genes to adherence of *Escherichia coli* O157:H7 to epithelial cells. Iowa state University,
- Orden, J.A., Cortes, C., Ruiz-Santa-Quiteria, J.A., Martinez, S., de la Fuente, R., 2005. Detection of the *saa* gene in verotoxin-producing *Escherichia coli* from ruminants. *J. Vet. Diag. Invest.* 17, 65-67.
- Passalacqua, K.D., Charbonneau, M., O'Riordan, M.X.D. 2016. Bacterial metabolism shapes the host-pathogen interface. In *Virulence mechanisms of bacterial pathogens*, Kudva, I., ed. (Washington, DC, ASM).
- Puttamreddy, S., Cornick, N.A., Minion, F.C., 2010. Genome-wide transposon mutagenesis reveals a role for the pO157 genes in biofilm development in *Escherichia coli* O157:H7 EDL933. *Infect. Imm.* 78, 2377-2384.
- Rice, D.H., Sheng, H.Q., Wynia, S.A., Hovde, C.J., 2003. Rectoanal mucosal swab culture is more sensitive than fecal culture and distinguishes *Escherichia coli* O157:H7 colonized cattle and those transiently shedding the same organism. *J. Clin. Microbiol.* 41, 4924-4929.
- Roe, A.J., Tysall, L., Dransfield, T., Wang, D., Fraser-Pitt, D., Mahajan, A., Constandinou, C., Inglis, N., Downing, A., Talbot, R., Smith, D.G.E., Gally, D.L., 2007. Analysis of the expression, regulation and export of NleA-E in *Escherichia coli* O157:H7. *Microbiol.* 153, 1350-1360.
- Schlager, T.A., Hendley, J.O., Bell, A.L., Whittam, T.S., 2002. Clonal diversity of *Escherichia coli* colonizing stools and urinary tracts of young girls. *Infect. Imm.* 70, 1225-1229.
- Shell, S.S., Prestwich, E.G., Baek, S.H., Shah, R.R., Sassetti, C.M., Dedon, P.C., Fortune, S.M., 2012. DNA Methylation Impacts Gene Expression and Ensures Hypoxic Survival of *Mycobacterium tuberculosis*. *PLoS Pathogens* 9, e1003419.
- Van Diemen, P., Dziva, F., Stevens, M.P., Wallis, T., 2005. Identification of enterohemorrhagic *Escherichia coli* O26:H genes required for intestinal colonization in calves. *Infect. Imm.* 73, 1735-1743.

- Vlisidou, I., Marches, O., Dziva, F., Mundy, R., Frankel, G., Stevens, M.P., 2006. Identification and characterization of EspK, a type III secreted effector protein of enterohemorrhagic *Escherichia coli* O157:H7. FEMS Microbiol. Lett. 263, 32-40.
- Wells, J.E., Shackelford, S.D., Berry, E.D., Kalchayanand, N., Guerini, M.N., Varel, V.H., Arthur, T.M., Bosilevac, J.M., Freetly, H.C., Wheeler, T.L., Ferrell, C.L., Koohmaraie, M., 2009. Prevalence and level of *Escherichia coli* O157:H7 in feces and on hides of feedlot steers fed diets with or without wet distillers grains with solubles. J. Food Protect. 72, 1624-1633.
- Whittam, T.S., 1989. Clonal dynamics of *Escherichia coli* in its natural habitat. Antonie van Leeuwenhoek 55, 23-32.

Table 1. Genes identified by STM as being required for persistent colonization of sheep by *E. coli* O157:H7

EDL gene	Gene name	Protein	Function	Site directed mutant made	Mutant tested in vivo
Z0065		Hypothetical protein	Unknown	No	NA
Z1745	<i>thiK</i>	Thiamine kinase	Thiamine metabolism	Yes	Yes
Z2164	<i>ftrA</i>	ftrA transcriptional activator	Regulator, carbon & amino acid metabolism	Yes	Yes
Z2389		Putative DNA methyltransferase (on a prophage)	Cysteine & methionine metabolism	Yes	Yes
Z3199	<i>wbdP</i>	WbdP glycosyl transferase	Surface polysaccharide metabolism	Yes	Yes
Z3491	<i>nrdB</i>	Ribonucleoside diphosphate reductase 1 β -subunit	Purine & pyrimidine metabolism	No	NA
Z4328	<i>nleB</i>	Hypothetical protein	Non-LEE-encoded effector protein	Yes	No
Z4995	<i>bax</i>	Hypothetical protein	Unknown function	Yes	No
Z5393	<i>yihF</i>	Hypothetical protein	Unknown function	Yes	No

Table 2. Specific growth rates of mutants grown alone and in a one-to-one ratio with the parent strain

EDL gene	Gene name	Specific growth rate alone ^a	Specific growth rate in presence of parent strain	Mutant alone vs with parent ^b	Mutant vs parent strain ^c
Z1745	<i>thiK</i>	6.19	6.30	0.48	0.2
Z2164	<i>ftrA</i>	7.05	6.65	0.003	0.46
Z2389		6.55	5.88	0.12	0.07
Z3199	<i>wbdP</i>	6.31	5.99	0.48	0.03
Z4328	<i>nleB</i>	6.83	6.44	0.12	0.51
Z4995	<i>bax</i>	6.16	6.79	0.0002	0.02
Z5393	<i>yihF</i>	7.08	6.64	0.05	0.19
Parent		6.89			

^athe specific growth rate was calculated using the formula $k=(t2-t1)/2.303$

^bt-test was done to compare the specific growth rate of the mutant alone and when grown competitively with the parent strain

^ct-test was done to compare the specific growth rate of the mutant with that of the parent strain

Figure

Figure 1. Schematic representation of the STM selection process that was used in this study. A total of 36 sheep were used for the first round of screen and 10 sheep were used in the second round of screening.

Fig 2. Binding of *E. coli* O157:H7 mutants to Hep-2 cells.

SEM, standard error of the mean of % adherence. *ΔehxD* mutant was used as a negative control

(Puttamreddy et al., 2010). a, b denote significant differences between strains $P < 0.05$, paired t-test

Figure 3. Fecal shedding of the $\Delta wbdP$ mutant and wild type *E. coli* O157:H7 recovered from sheep. Circles represent individual animals (● wild type; ○ $\Delta wbdP$) and the bars represent the mean (- wild type; + $\Delta wbdP$).

Figure

1 Figure 4. Fecal shedding of 3 additional mutants and wild type strain 86-24 of *E. coli* O157:H7. These mutants were
2 inoculated into 2 sheep each. Symbols represent individual animals (● wild type; □ Δ *ftrA*; △ Δ Z2389; ◇ Δ *thiK*), the bar
3 represents the mean of the wild type strain.

4
5

6
7
8
9
10