

## Genotypic Diversity of *Escherichia coli* in a Dairy Farm

Insook Son, Jo Ann S. Van Kessel, and Jeffrey S. Karns

### Abstract

Dairy cattle are known reservoirs of pathogenic *Escherichia coli*, but little is known about the dynamics of *E. coli* in dairy cows or within the dairy farm environment. This study was conducted to evaluate the diversity and distribution of *E. coli* strains in a dairy farm using pulsed-field gel electrophoresis and to determine the relationships between *E. coli* isolated from feces and throughout the farm environment. Water from watering troughs, feces from cows, manure composites, milk, and milk filters were collected on December 2005 and December 2006. Isolates were analyzed by PCR for phylogenetic grouping (A, B1, B2, and D) and for the presence of virulence genes associated with enteropathogenic *E. coli* and enterohemorrhagic *E. coli* strains. Most of the isolates were in groups A (22%) and B1 (64%), while 4% and 11% of the isolates were within groups B2 and D, respectively. Enterohemorrhagic *E. coli* and enteropathogenic *E. coli* virulence genes were detected in strains from the feces of three cows and in one manure composite, and *E. coli* O157:H7 was present in one manure composite. Pulsed-field gel electrophoresis analysis resulted in 155 unique restriction digestion patterns (RDPs) among 570 isolates. *E. coli* isolates from water, manure composites, feces, milk, and milk filters grouped into 34, 65, 76, 4, and 6 clusters (identical RDPs), respectively. There was little diversity of isolates within individual fecal samples; however, high diversity was observed between fecal samples. Diversity was high within the water and composite samples. Some RDPs were common to multiple sample types. Although there were common RDPs between the 2005 and 2006 samplings, the *E. coli* populations were quite distinct between these two sampling times. These results demonstrate a high degree of diversity for *E. coli* within a dairy farm and that assigning a single environmental isolate to a particular farming operation would require the testing of an impractical number of isolates.

### Introduction

*ESCHERICHIA COLI* ARE PART of the complex microbial community in the gastrointestinal tracts of healthy mammals. Although some strains of *E. coli* are pathogenic, the degree of pathogenic potential varies between strains and pathogenic forms (Thorpe *et al.*, 2002) represent a small proportion of the species. Human exposure to pathogenic *E. coli* such as enterohemorrhagic strains is typically through contaminated food and water or direct contact with infected animals. Recent increases in human illness due to pathogenic *E. coli* have raised considerable interest in the diversity and epidemiology of *E. coli*, especially of strains associated with animal production. Contamination of food and water with animal feces can lead to exposure of humans to several different zoonotic pathogens, and because of its ubiquity in feces, the presence of *E. coli* is often considered an indication of fecal

contamination for water and food safety purposes (EPA Office of Science and Technology, 2000).

Dairy cattle are known reservoirs of both pathogenic and nonpathogenic *E. coli*, but little is known about the dynamics of *E. coli* in dairy cows or within the dairy farm environment (Wells *et al.*, 1991; Whipp *et al.*, 1994). Dairy farms are complex environments. Animals are grouped and housed in various locations depending on age and stage of lactation. Dietary components are stored in large quantities and typically include crops that are grown on the farm as well as purchased supplements such as grains and waste products from food factories. Individual animal groups are fed different diets, and water troughs are located throughout the animal housing areas. Milking cows are moved at least twice per day through common alleys to a milking parlor, and manure is typically moved daily to a storage unit and then spread on the fields at a later date. A typical dairy cow can excrete 40 to 50 kg of feces

daily (Wilkerson *et al.*, 1997), and although biocontrol strategies are often in place, it is very challenging to control the spread of pathogens throughout the farm.

Because *E. coli* is ubiquitous in feces, phenotypic and genotypic typing of *E. coli* has been proposed as a means of tracking sources of fecal contamination. Several studies that were focused on the characterization of enterohemorrhagic *E. coli* O157:H7 found that comparison of pulsed-field gel electrophoresis (PFGE) profiles allowed discrimination between isolates within the serotype (Faith *et al.*, 1996; Shere *et al.*, 1998; Welinder-Olsson *et al.*, 2002; Lahti *et al.*, 2003; Avery *et al.*, 2004; Arthur *et al.*, 2007). When Faith *et al.* (1996) and Lahti *et al.* (2003) looked at multiple *E. coli* O157:H7 isolates associated with individual farms, they found minimal genotypic variation within an operation. Some studies have demonstrated high genomic variability among commensal, nonpathogenic *E. coli* strains from healthy pigs, cattle, and humans (Jarvis *et al.*, 2000; Duriez *et al.*, 2001; Leung *et al.*, 2004; Higgins *et al.*, 2007; Houser *et al.*, 2008). Houser *et al.* (2008) compared 100 isolates from 10 healthy dairy cows and, based on a combination of PFGE analysis, serotyping, phylogenetic grouping, and antimicrobial resistance testing, concluded that there was a significant level of phenotypic and genetic diversity in *E. coli* isolates from individual animals and between animals. Thus, although molecular techniques like PFGE may be useful for tracking specific pathogenic strains of *E. coli*, a high degree of diversity among the total *E. coli* might preclude their use for tracking sources of environmental contamination.

This study was conducted to evaluate the diversity and distribution of *E. coli* strains in a dairy farm, to evaluate the change in the population over time, and to determine the relationships between *E. coli* isolated from feces and throughout the farm environment.

## Materials and Methods

### Sample collection

The herd and the logistics of herd sampling were described previously (Van Kessel *et al.*, 2007). Briefly, the herd consisted of approximately 105 mature cows that were housed in a three-row free stall barn and milked twice daily. The cows were divided into three groups (group 1, group 2, and dry cow), and each was housed in a separate section of the barn. Cows from group 1 and group 2 were kept in a shared holding pen while awaiting milking, and utilized a common alley to return to the barn. Alleys were scraped daily and manure from group 1 and group 2 was pushed through the dry cow pen alleys during the cleaning process. Heifers were raised at a contract facility geographically separated from the farm, and replacement animals were occasionally purchased. The farm was located in proximity to other dairy farms.

Water, feces, manure composites, milk, and milk filters used in this study were collected in December 2005 and December 2006. Fecal samples were collected directly from the rectum and placed into sterile vials using individual disposable sleeves. For this study, 15 cows were chosen randomly at each time point. Manure composite samples were taken from the group 1, group 2, and dry cow alleys, the manure storage pit, and the postweaned calf pen. To obtain a representative sample, grab samples were mixed, and representative aliquots were placed in Whirl-pak® bags (Nasco, Ft. Atkinson, WI). Trough water samples (500 mL to 1 L) were obtained by

mixing the water and submersing a sterile bottle in the trough. Water samples were collected at four different sampling sites: the dry cow, pen, two locations in the group 1 pen (group 1 middle and group 1 far end) and the group 2 pen. All samples were packed on ice and transported overnight to the laboratory. Samples were processed within 24 h of collection.

### Bacteriological analysis

For fecal and manure composite samples, approximately 25 g of samples was weighed into a filtered stomacher bag (GSI Creos, Tokyo, Japan). A 2:1 (w/w) dilution of 1% buffered peptone water (BPW) was added to the bag and pummeled in an automatic bag mixer (BagMixer®; Interscience Laboratories, Weymouth, MA) for 2 min. For enrichment of *E. coli*, 5 mL of filtrate was added to 5 mL of double-strength EC broth (Difco, Detroit, MI). Water samples were filtered (100 mL) through sterile 0.45  $\mu$ m cellulose filters (47 mm; Osmonics, Westborough, MA). Filters were placed in a 15 mL tube containing 5 mL of double-strength EC broth and 5 mL of BPW and vortexed. For bulk tank milk, 5 mL of milk was mixed with 5 mL of double-strength EC broth. The milk filter was cut into small pieces, mixed with 1% BPW (2:1, w:w) in a stomacher bag, and pummeled for 2-min periods. Then 5 mL of liquid was added to 5 mL of double-strength EC broth. Enrichments were incubated at 37°C for 18–24 h. Two milliliters of the enrichment was centrifuged (13,000 g, 2 min), the supernatant removed, and the resulting pelleted material was suspended in 0.5 mL of a 1:1 mixture of the 2 $\times$  freezing medium described by Schleif and Wensink (1981) and Lennox broth (Gibco Laboratories, Long Island, NY) and frozen at –80°C. To re-culture a sample, frozen biomass was added to 200  $\mu$ L of Lennox broth and incubated at 37°C for 2 h. Fifty-microliter aliquots of the enrichments were plated onto MacConkey agar (Remel, Lenexa, KS) and incubated at 37°C overnight. Up to 20 randomly chosen and well-isolated colonies with pink color and typical *E. coli* morphology were patched onto Simmons citrate (BD Diagnostics, Sparks, MD), MacConkey (Remel), MacConkey with 4-methylumbelliferyl beta-D-glucuronide (MUG) (Remel), and L-agar (Lennox broth base with 1.5% agar; Gibco Laboratories). Simmons citrate, MacConkey, and L-agar plates were incubated at 37°C overnight. MUG agar plates were incubated at 37°C for 5 h and then checked for the fluorescence using a UV viewing cabinet (IDEXX Laboratories, Westbrook, Maine) to see the presence of  $\beta$ -D-glucuronidase activity. Isolates with the inability to ferment citrate, ability to ferment lactose, and the presence of  $\beta$ -D-glucuronidase activity were considered as phenotypically *E. coli*. Fifteen isolates were picked for each composite sample (Table 1), and 10 isolates were picked for each of the remaining sample types. In total, 570 isolates were chosen for this study.

For PCR analysis, DNA was extracted from a single isolated colony using 200  $\mu$ L InstaGene Matrix (Bio-Rad Laboratories, Hercules, CA) following the manufacturer's directions. The DNA preparations were stored at –20°C and were analyzed for the *gadAB* and phylogenetic grouping by conventional PCR. Presumptive *E. coli* isolates that were  $\beta$ -D-glucuronidase-negative were tested for the presence of the *E. coli gadAB* gene region to confirm their identity as *E. coli* by the PCR method of McDaniels *et al.* (1996). Each 50  $\mu$ L reaction mixture contained 1.5 mM MgCl<sub>2</sub> (Roche, Indianapolis, IN),

TABLE 1. DISTRIBUTION OF *ESCHERICHIA COLI* ISOLATES COLLECTED FROM VARIOUS SOURCES ON A DAIRY FARM AT EACH OF TWO COLLECTIONS (DECEMBER 2005 AND DECEMBER 2006)

Sample source	No. of samples per collection	No. of isolates per sample	Total no. of isolates
Feces	15	10	300
Manure composites <sup>a</sup>	5	15	150
Trough water <sup>b</sup>	4	10	80
Bulk milk	1	10	20
Milk filter	1	10	20
Total			570

<sup>a</sup>Samples were taken from the group 1, group 2, and dry cow alleys, manure storage pit, and the postweaned calf pen.

<sup>b</sup>Trough water samples were taken from the dry cow pen, two locations in the group 1 pen (middle and far end), and the group 2 pen.

2.5 mM each deoxyribonucleoside triphosphate (Takara Bio, Shiga, Japan), 50 pmol of each primer, 1 U of Taq DNA polymerase (Invitrogen, Carlsbad, CA), and 5  $\mu$ L of 10 $\times$  buffer (supplied with Taq polymerase). PCR conditions were denaturation for 2 min at 94°C, 30 cycles of 15 sec at 94°C, 1 min at 58°C, 1 min at 72°C, and a final extension step of 7 min at 72°C. Amplification products were separated by electrophoresis in a 1% agarose gel containing 10  $\mu$ g/mL ethidium bromide with 1 $\times$  Tris-borate-EDTA buffer (Sigma, St. Louis, MO) and examined by UV transillumination using the Gel Doc system (Bio-Rad Laboratories). A molecular weight standard (Roche) with a 100-bp DNA ladder was included on each gel.

#### Phylogenetic grouping

A triplex PCR was performed using phylogenetic group-specific primers for two genes (*chuA* and *yjaA*) and an anonymous DNA fragment (TspE4C2) according to the method of Clermont *et al.* (2000) except that the Taq polymerase and buffer from Invitrogen were used. Amplification products were analyzed by agarose gel electrophoresis (2.5% agarose, 0.5 $\times$  Tris-borate-EDTA buffer) and visualized as described above.

#### Virulence gene analysis

The methods for the detection of Shiga toxin 1 (*stx1*) and Shiga toxin 2 (*stx2*) genes, *eaeA*, and  $\gamma$ -*tir* associated with enterohemorrhagic *E. coli* to determine the incidence of possible enteropathogenic or enterohemorrhagic forms of *E. coli* were performed as described by Karns *et al.* (2007). For confirmation, agglutination tests on presumptive isolates of O157:H7 were performed using a RIM™ *E. coli* O157:H7 latex test kit (Remel) following manufacturer's instructions.

#### Pulsed-field gel electrophoresis

All *E. coli* isolates ( $n=570$ ) were subtyped using the 24 h standardized PFGE protocol for *E. coli* developed by the Centers for Disease Control and Prevention (CDC) and PulseNet (Ribot *et al.*, 2006) with *XbaI* as the restriction enzyme. *Salmonella enterica* serotype Braenderup H9812 digested with *XbaI* was used as the molecular size standard. All isolates

were grown on tryptic soy agar containing 0.01% yeast extract at 37°C for 24 h for use in preparing agarose plugs. The restriction fragments of the *E. coli* isolates were separated on 1% gels using a CHEF-DR® III (Bio-Rad Laboratories). The gels were stained for 30 min with ethidium bromide (10  $\mu$ g/mL; Sigma), destained in distilled water, and visualized with a UV gel documentation system (Bio-Rad Laboratories).

#### Cluster analysis and group statistics

The PFGE patterns of *E. coli* were analyzed using BioNumerics software (Version 5.0; Applied Maths, Austin, TX). The optimization setting was 0.5%, and band position tolerance was 1.5%. Suspected double bands were checked by examining the plotted densitometric curves of the PFGE profiles. Bands greater than 1135 kb or less than 20.5 kb were excluded from PFGE analysis since they fell outside the coverage of the standard. Cluster analysis was performed using the Dice coefficient and unweighted pair group method using arithmetic averages. A similarity score value of 95% was used as the cutoff for designating strain types.

Simpson's index of diversity (SID) as described by Hunter and Gaston (1988) was used to assess the suitability of PFGE for differentiation of *E. coli* isolates and to describe the diversity of *E. coli* subtypes within and between different sample populations.

Jackknife analysis was used to determine how accurately isolates could be assigned to host groups. This was done using the maximum similarity and the "spread equally" option in the BioNumerics software. For each group to be assigned, a similarity value is calculated between each sample and every other sample in its own group using the Dice similarity coefficient. Each sample is taken out of its own group and placed in another group, and similarities are again calculated. If any entry has a higher similarity value with another group, this is considered a violation of that group. The percentage of entries having the highest similarity score with their own group is shown in the diagonal of the jackknife matrix.

## Results

#### Cluster analysis

A total of 570 *E. coli* isolates were isolated from feces, manure composites, trough water, bulk tank milk, and milk filter samples that were collected on a single dairy farm in December 2005 and December 2006. The sample sources, number of samples from each source, and the number of *E. coli* isolated from each type of sample are detailed in Table 1. When *XbaI*-digested genomic DNA from these isolates was analyzed by PFGE, a wide range of restriction digestion patterns (RDP) was observed. A 95% similarity cutoff cluster analysis indicated that there were a total of 155 unique clusters of RDPs within the complete group of 570 *E. coli* isolates (Fig. 1). Only 2 clusters represented more than 20 isolates, while 13 clusters contained between 10 and 19 isolates each. The remaining 140 clusters contained less than 10 isolates each, and 66 of these represented only 1 isolate.

The number of unique RDPs within a sampling type was variable at each sampling time (Table 2). The *E. coli* isolates from the manure composites were the most diverse sample group; however, diversity was also high in the water and the fecal populations. The fecal isolates were relatively homogeneous within an animal, but there was little overlap between

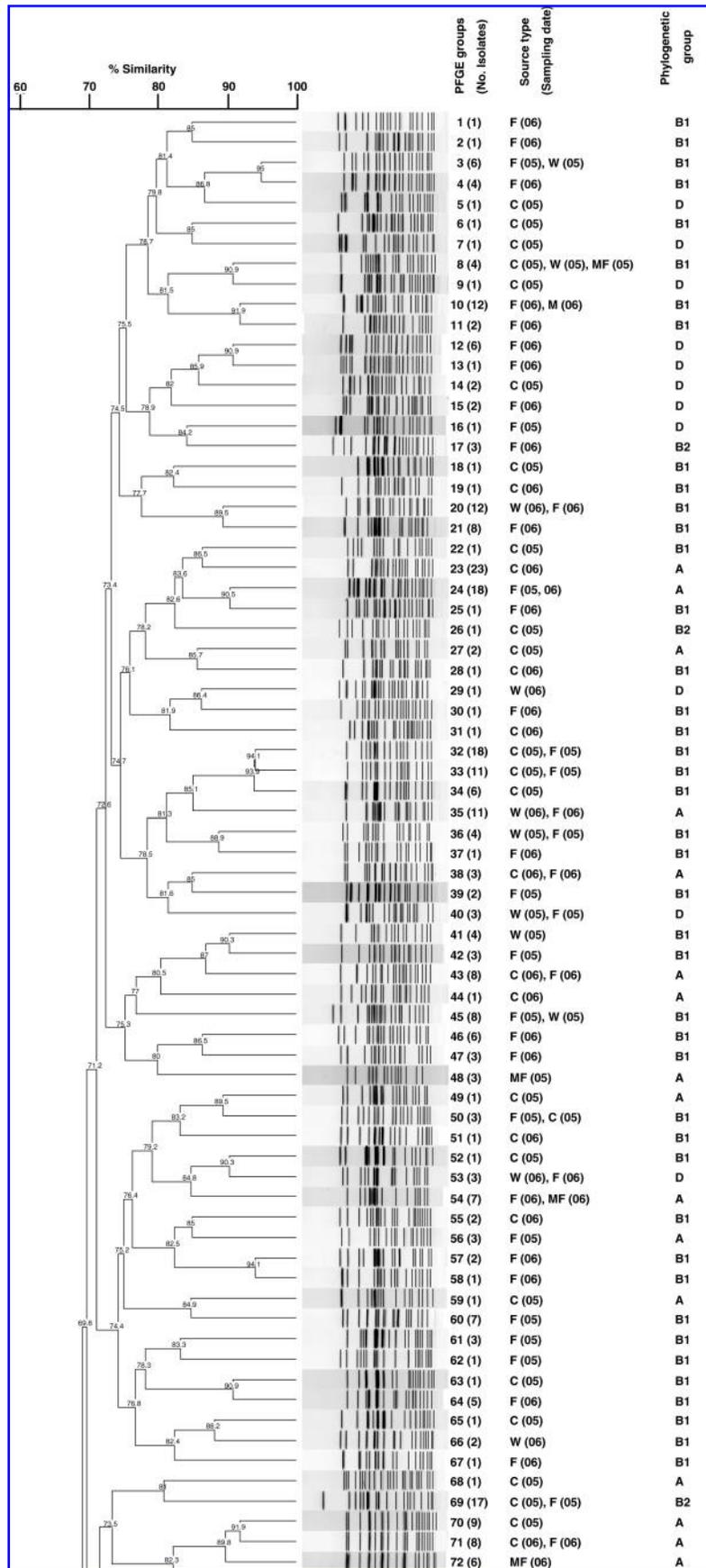


FIG. 1. Pulsed-field gel electrophoresis pattern clusters (155) of *Escherichia coli* isolates ( $n = 570$ ) from milk, milk filters, trough water, feces, and manure composites collected in December 2005 and December 2006. A single isolate was chosen to represent each cluster. Sampling site: W, trough water; C, manure composite; F, feces; M, milk; MF, milk filter. Sampling dates: December 5, 2005, and December 5, 2006.

(continued)

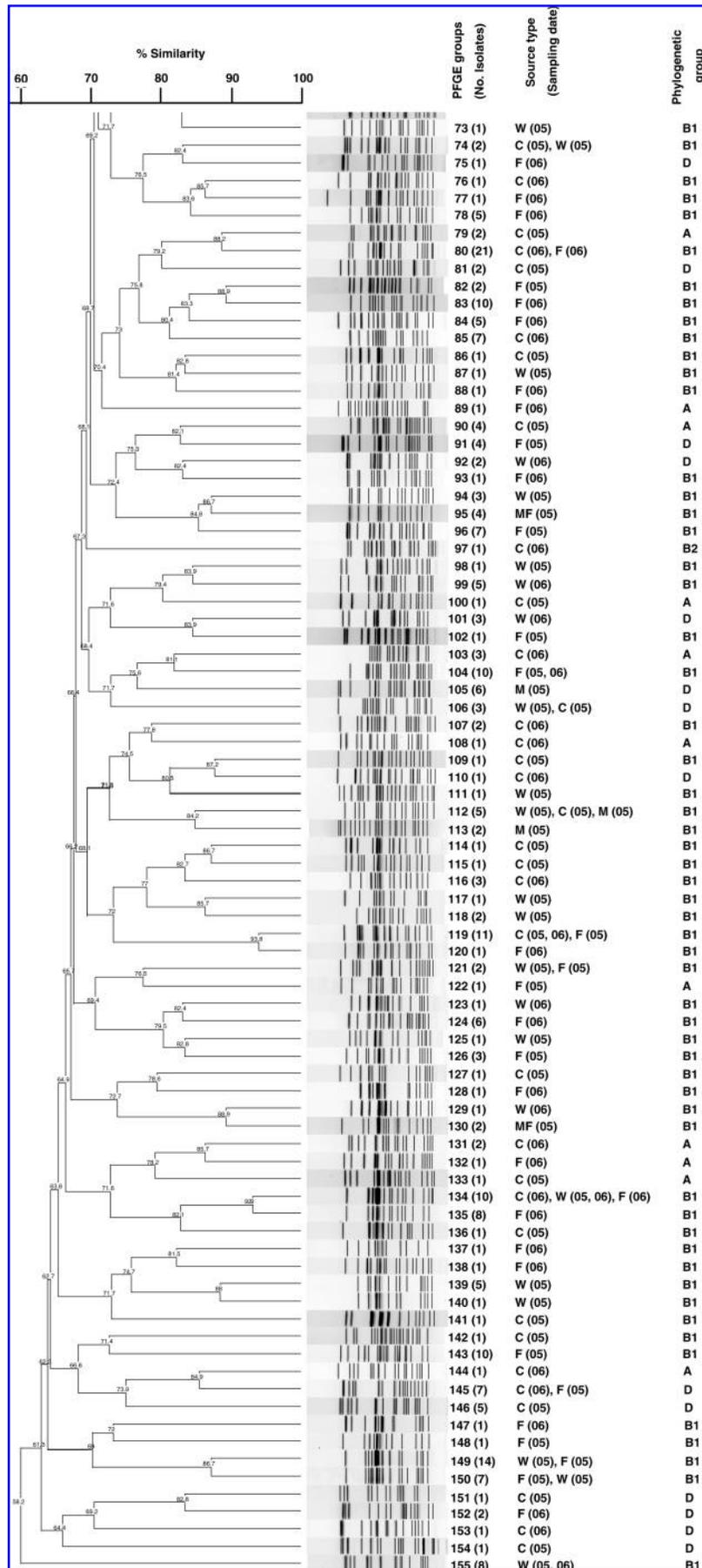


FIG. 1. (Continued).

TABLE 2. THE NUMBER OF RESTRICTION DIGESTION PATTERNS (RDPs) AND SIMPSON'S INDEX OF DIVERSITY (SID) OF *ESCHERICHIA COLI* ISOLATES WITHIN SAMPLING GROUPS AT TWO COLLECTION TIMES

Sampling year	Sample Source											
	Composites		Fecal		Trough water		Milk		Milk filter		Whole	
	RDPs	SID	RDPs	SID	RDPs	SID	RDPs	SID	RDPs	SID	RDPs	SID
2005	41	0.96	31	0.95	24	0.97	3	0.62	4	0.78	85	0.98
2006	25	0.88	47	0.97	12	0.89	1	0.00 <sup>a</sup>	2	0.53	76	0.97
2005 + 2006	65	0.96	76	0.98	34	0.96	4	0.67	6	0.84	155	0.99

<sup>a</sup>All isolates had the same RDP. Therefore, the SID could not be calculated and was assumed to be 0.00.

animals, and thus the total fecal *E. coli* population was very diverse. In contrast to the water, fecal, and composite samples, little diversity was observed in the *E. coli* strains that were isolated from milk. There was somewhat more diversity in the milk filter isolates; however, there were no shared isolates between milk and milk filter samples. The *E. coli* populations were unique for each year of sampling, and there was little or no temporal overlap in all sampling groups. The composite manure, water, milk, and milk filter had more RDP types in 2005 than in 2006, while the number of RDP types was higher in fecal samples in 2006.

Isolates from the manure composite samples (150) were grouped into 65 unique RDPs (Fig. 1). In 2005 the 75 isolates from manure composites were distributed in 41 clusters. The 15 isolates from the area of the barn that housed group 1 were represented by 12 distinct RDPs, while the 15 isolates from the group 2 composite were grouped into eight clusters. Only one cluster contained isolates from both groups 1 and 2. There was less diversity in the manure from the floor of the area of the barn that housed the dry cows, with 15 isolates grouped into only two clusters. Each of the 15 isolates from the manure storage pit yielded a unique RDP, none of which was found in groups 1, 2, or the dry cow samples despite the fact that the manure from those areas is pushed directly into the pit several times a day. The 15 isolates from a separate area of the farm that housed postweaned calves were grouped into five clusters with one cluster of 2 isolates characterized as pathogenic O157:H7. Less diversity was observed in composite manures from 2006 with the 75 isolates grouped into only 25 clusters. Isolates from groups 1 and 2, and from the dry cows grouped into three, five, and nine clusters, respectively. The isolates from the pit grouped into nine clusters, five of which shared RDPs with isolates from group 1, group 2, or the dry cow composites. The 15 isolates from the postweaned calf area grouped into six clusters, one of which shared RDPs with isolates from the dry cow composite sample. No RDPs were shared between groups 1, 2, or the dry cow composites in 2006. There was little temporal overlap with only one RDP (three isolates) represented in manure composites from both sampling times.

Diversity was also high in the water samples with 34 RDPs represented by the 80 isolates. The main area where the milking and dry cows are held has four water troughs, two in the group 1 area, with one at each end (far end and middle), one in the group 2 area, in proximity to the middle trough of the group 1 pen, and one in the dry cow area. Although access to a particular trough is generally limited to the cows in the particular group it is intended to service, other cows may access water troughs outside their areas as they are moved

around the barn for milking or for alley cleaning. For the most part there was little overlap in the strains of *E. coli* isolated from water troughs. In 2005 one RDP was found in both the group 1 middle and the group 2 troughs, which, as noted above, are in close physical proximity. In 2006, identical isolates were again found in the group 1 middle and group 2 troughs; however, at this sampling a substantial number of identical isolates were found in the group 1 middle and group 1 far end troughs both of which are readily accessible to all the cows in group 1. Forty-four of the 80 isolates obtained from both sets of water samples were unique to the water troughs and grouped into 18 clusters, while 25 isolates (10 RDPs) were found both in water and feces, 5 (3 RDPs) in water and manure composites, and 5 isolates (1 RDP) found in water, manure composites, and feces. As with the manure composites, there was little temporal overlap in water samples with only two RDPs (13 isolates) found at both sampling times.

As a group, fecal samples taken from individual cows displayed a high degree of diversity. In 2005 there were 31 unique RDPs among the 150 isolates obtained from 15 cows. While the ratio of RDPs to samples is lower than in composites or water, only five RDP clusters shared isolates from two or more cows. A similar result was obtained with samples from 2006 with 47 RDPs representing 150 isolates but with only 6 RDPs shared by isolates from multiple cows. Thus, although diversity for the group was high, the diversity within individual cows was relatively low. As with manure composites and water samples, there was little temporal overlap with only two RDPs representing 28 isolates found in both 2005 and 2006. Out of 76 unique RDPs found in fecal samples over the 1-year period, only 11 were also found in manure composites.

There was low diversity among *E. coli* isolated from each of the milk samples collected in 2005 and 2006. In 2005 there were three unique RDPs from 10 isolates, while in 2006 all 10 isolates yielded identical RDPs. One cluster that contained milk isolates from 2005 also contained isolates from water and composite manure samples collected on the same date, while milk isolates from 2006 shared one RDP with fecal isolates from 2006. In both years the *E. coli* from milk filter samples were somewhat more diverse than the milk isolates with the 2005 isolates grouped into four clusters and those from 2006 grouped into two clusters. In 2005 one cluster containing milk filter isolates also contained isolates from water and composite manure from the group 1 area. In 2006 one of the clusters also contained fecal isolates. No RDPs were shared between milk and milk filters at either sampling and there was no temporal overlap.

To further assess the diversity of the *E. coli* isolates, SID values were calculated for both the individual and combined

sample groups (Table 2). The SID for the composite, fecal, and water groups were all 0.9 or higher at each sampling time and for the combined sampling times. The lowest SID was for the milk filter in 2006 (0.53), but the SID values for milk were also low (0.67 for 2005 and 2006 combined). To assess the variation within fecal samples, SID values were also calculated for each individual sample (Table 3). The 10 isolates within each of four fecal samples (two for each sampling time) had indistinguishable RDPs indicating no diversity. In contrast, the 10 isolates from one fecal sample that was collected in 2006 had nine different RDPs, and the SID was calculated to be 0.98. For the most part, however, the SID values were less than 0.8 in the remainder of the fecal samples and the average SID was 0.46 for 2005 and 0.54 for 2006. When the SID was calculated for individual composite and water samples, it appeared that these samples in general contained a more diverse population of *E. coli* than the individual fecal samples. The SID values for the 10 composite samples were calculated to be between 0.42 and 1, while the SID values for the 8 water samples were between 0.56 and 0.96.

#### Jackknife analysis

Jackknife analyses using Dice similarity was conducted on the *E. coli* isolates using five defined groups based on isolate

origin. Initially, isolates from the two sampling years were separated for analysis (Table 4) and then a jackknife analysis was conducted on combined isolates from both years. The average rates of correct group classification ranged from 63.8% to 90.0% and 50.0% to 81.7% for 2005 and 2006, respectively, and 67.4 to 85.1 when the isolates from the 2 years were combined.

For individual and combined sampling years, the fecal group had a higher percentage of isolates with similarity scores defining them as belonging to the group from which they were isolated rather than the composite and water groups. However, in the combined year analysis, 7.6% and 15% of the fecal isolates had higher similarity scores with the water and composite isolates, respectively. The analysis suggests that an isolate from water had a 23.8% chance of being identified as a fecal isolate in 2005 (Table 4). The milk filter group in 2006 had the highest percentage of isolates for which the similarity scores defined them as belonging to other groups, although this group consisted of only 10 isolates.

When the isolates were grouped based on sampling years, the 2005 group was determined to have a similarity score of 87.2%, and 12.8% of the isolates were more similar to the isolates from the 2006 group (Table 5). Similarly, in the 2006 group, more than 90% of the isolates were grouped within the 2006 group, and only 8.6% of the 2006 isolates were more closely related to the 2005 isolates versus the 2006 isolates.

TABLE 3. THE NUMBER OF RESTRICTION DIGESTION PATTERNS (RDPs) AND SIMPSON'S INDEX OF DIVERSITY (SID) OF *E. COLI* IN FECAL SAMPLES AT TWO COLLECTION TIMES

Sampling date	Cow IDs	RDPs	SID
December 2005	82	5	0.82
	99	3	0.51
	107	2	0.47
	124	2	0.47
	126	3	0.62
	131	4	0.71
	215	3	0.62
	227	2	0.47
	410	1	0.00 <sup>a</sup>
	429	2	0.20
	455	2	0.47
	464	2	0.47
	478	3	0.38
	483	1	0.00
487	3	0.64	
December 2006	103	4	0.64
	424	9	0.98
	432	3	0.60
	446	7	0.91
	450	2	0.36
	475	1	0.00
	477	3	0.38
	486	3	0.51
	491	3	0.51
	504	4	0.71
	511	2	0.36
	519	1	0.00
	524	5	0.76
	525	5	0.76
530	4	0.64	

<sup>a</sup>All isolates had the same RDP. Therefore, the SID could not be calculated but was assumed to be 0.00.

#### Phylogenetic grouping

Isolates were classified into four phylogenetic groups when triplex PCR was used to identify the presence or absence of *chuA*, *yjaA*, and TspE4C2 (Table 6). In 2005 68% of the isolates (194) were in group B1, and this was the most prevalent group. Groups A, B2, and D contained 13%, 6%, and 13% of the isolates, respectively. The most common group in 2006 was also B1 (60%), but the percentage of group A was higher (31%) than in 2005. When the results from the 2 years were combined, group B1 was the most prevalent group (64%) followed by group A (22%), group D (11%), and group B2 (4%).

#### Virulence gene analysis

When the *E. coli* isolates were analyzed for the presence of the intimin and shiga toxin genes, the intimin gene (*eaeA*) was detected in four isolates from one postweaned calf composite sample (Table 7). Subsequent analyses indicated that two of these isolates also contained *stx2* and the  $\gamma$ -allele of the translocated intimin receptor ( $\gamma$ -*tir*). Phylogenetic grouping PCR determined that the latter two isolates fell within group A, while the isolates containing all three virulence genes were classified as group D. Additionally, five *E. coli* isolates from fecal samples contained one or both shiga-like toxin genes (*stx1* and *stx2*), and these isolates were classified into phylogenetic groups B1 or D. The remaining 561 isolates were negative for all three virulence genes.

#### Discussion

Tracking bacterial transfer or dissemination in complex environments such as animal production facilities is important for developing and evaluating strategies that mitigate pathogen persistence and transfer within the farm and to the surrounding environment. DNA-based methodologies have

TABLE 4. ASSIGNMENT OF ISOLATES TO SAMPLING GROUPS BY USING PULSED-FIELD GEL ELECTROPHORESIS AND JACKKNIFE ANALYSIS

Assigned group	2005				
	% of <i>E. coli</i> isolates in assigned group <sup>a</sup>				
	Feces	Composite	Water	Milk	Milk filter
Feces	<b>74.0</b>	19.8	23.8	0.0	0.0
Composite	17.0	<b>63.8</b>	12.9	6.7	5.0
Water	9.0	14.4	<b>60.4</b>	6.7	5.0
Milk	0.0	0.7	1.7	<b>86.7</b>	0.0
Milk filter	0.0	1.3	1.3	0.0	<b>90.0</b>
Assigned group	2006				
	% of <i>E. coli</i> isolates in assigned group				
	Feces	Composite	Water	Milk	Milk filter
Feces	<b>81.7</b>	23.3	26.7	20.0	50.0
Composite	11.3	<b>72.0</b>	1.7	0.0	0.0
Water	5.1	4.0	<b>69.2</b>	0.0	0.0
Milk	1.0	0.7	2.5	<b>80.0</b>	0.0
Milk filter	0.9	0.0	0.0	0.0	<b>50.0</b>
Assigned group	2005 + 2006				
	% of <i>E. coli</i> isolates in assigned group				
	Feces	Composite	Water	Milk	Milk filter
Feces	<b>76.5</b>	23.9	23.3	25.0	10.0
Composite	15.0	<b>67.9</b>	6.7	3.3	2.5
Water	7.6	6.9	<b>67.3</b>	3.3	2.5
Milk	0.4	0.3	0.8	<b>68.3</b>	0
Milk filter	0.5	1.0	1.9	0.0	<b>85.0</b>

<sup>a</sup>Values in boldface on the diagonal indicate percentages of isolates correctly assigned to source groups.

been developed to differentiate bacterial isolates, and molecular typing is frequently used to identify sources of contamination or infection and to determine routes of transmission and persistence of bacterial strains within various environments (Lukinmaa *et al.*, 2004; Zadoks and Schukken, 2006). The method used must be sufficiently discriminatory to distinguish between various strains of the species in question, and PFGE is one method that is commonly used for discrimination between *E. coli* strains. Like other methods, PFGE is time consuming and costly. Data presented here suggest that molecular tracking of *E. coli* in dairy farm environments is not feasible due to the high level of strain diversity; the number of isolates necessary to describe the population is intractable.

Previous work by Houser *et al.* (2008) demonstrated a high level of diversity in the commensal *E. coli* population of healthy dairy cows. Based on both genotypic and phenotypic characterization of 100 isolates collected in one sampling time from 10 cows, they calculated that there were at least 412 strain types in this population. They concluded that at least 55 isolates were needed to describe the *E. coli* population in a single cow. The results presented here support the observations of high diversity (high SID) in the overall *E. coli* population in feces although the within-animal diversity (low SID) appeared to be somewhat less (Table 3) than the diversity described by Houser *et al.* (2008).

Even though the diversity of *E. coli* isolates within individual cows was relatively low in this study, the overall diversity for the fecal isolates was very high at each sampling time (SID >0.95). This observation is supported by the results obtained for manure composite and water samples where diversity was high both within and across samples. The fact that few RDPs (11 out of 130) were shared between composite and fecal samples indicates that the isolates from the environmental samples did not completely represent the commensal *E. coli* population in the herd and that many more manure composite isolates would be needed to adequately describe the *E. coli* population. Although there are other potential sources

TABLE 5. ASSIGNMENT OF ISOLATES TO COLLECTION YEAR BASED ON PULSED-FIELD GEL ELECTROPHORESIS AND JACKKNIFE ANALYSIS

Assigned group	Sampling year	
	2005	2006
2005	<b>87.2</b>	8.6
2006	12.8	<b>91.4</b>

Values in boldface indicate percentages of isolates correctly assigned to collection year.

TABLE 6. DISTRIBUTION OF *ESCHERICHIA COLI* STRAINS AMONG THE FOUR PHYLOGENETIC GROUPS<sup>a</sup>

Sampling date	Sample source	Phylogenetic groups				Total
		A	B1	B2	D	
December 2005	Water	0 (0)	37 (10%)	0 (0)	3 (5%)	40
	Manure composites	22 (18%)	36 (10%)	2 (9%)	15 (25%)	75
	Feces	11 (9%)	110 (31%)	16 (73%)	13 (21%)	150
	Milk	0 (0)	4 (1%)	0 (0)	6 (10%)	10
	Milk filter	3 (2%)	7 (2%)	0 (0)	0 (0)	10
	Total	36 (13%)	194 (68%)	18 (6%)	37 (13%)	285
December 2006	Water	9 (7%)	23 (6%)	0 (0)	8 (13%)	40
	Manure composites	37 (30%)	34 (9%)	1 (4%)	3 (5%)	75
	Feces	31 (25%)	103 (28%)	3 (14%)	13 (21%)	150
	Milk	0 (0)	10 (3%)	0 (0)	0 (0)	10
	Milk filter	10 (8%)	0 (0)	0 (0)	0 (0)	10
	Total	87 (31%)	170 (60%)	4 (1%)	24 (8%)	285
Total		123 (22%)	364 (64%)	22 (4%)	61 (11%)	570

<sup>a</sup>Data are given as numbers of strains (percentage in parentheses). The percentage is calculated as the number of strains from each source over the total number of strains belonging to each phylogenetic group × 100 in the first four columns and over the total number of strains × 100 in the last row.

of *E. coli* in the manure composite and water samples, animal feces is by far the overwhelming source. The effect of the enrichment culture used in this study is unknown, but would be expected to cause an underestimate rather than overestimate of diversity.

The *E. coli* population on this farm was apparently dynamic with few strains found at both sampling times. Based on the jackknife analysis, over 90% of the isolates were unique to the sampling year. It is possible that, had the diversity been more completely described, more temporal overlap would have been observed. Because 15 out of 105 cows were selected at each sampling, none of the cows that were sampled in 2005 were sampled again in 2006. Thus, no direct conclusions can be drawn on the stability of the population within individual cows. However, it is very plausible to suggest that the commensal *E. coli* populations within a cow would vary over time as the cow progresses through a lactation cycle with the associated changes in physiological stress and diet; transient *E. coli* have been described previously (Hancock *et al.*, 1998).

*E. coli* strains have been reported to fall into four main phylogenetic groups (A, B1, B2, and D) (Ochman and Selander, 1984; Herzer *et al.*, 1990). In this study, the majority of the *E. coli* isolates belonged to phylogenetic group B1. Recently, Higgins *et al.* (2007) reported that *E. coli* isolates of

bovine origin were predominately in groups B1 (44%) and A (38%). Houser *et al.* (2008) also found that the majority of the fecal *E. coli* isolates from healthy lactating cattle from one herd belonged to group B1.

Girardeau *et al.* (2005) characterized the distribution of virulence factors in 287 Shiga toxin-producing *E. coli* strains and determined that 70% of the isolates were in phylogenetic group B1 and 19% of the strains segregated into group A. While the majority of the Shiga toxin-producing *E. coli* and enterohemorrhagic *E. coli* classify as group B1, other strains can be found in groups A and E, the latter which is a minor group classified as D by the method used here (Escobar-Paramo *et al.*, 2004). The limited number of *stx*-containing isolates in this study were split equally between groups B1 and D. Additionally, the strain that contained the locus of enterocyte effacement, as evidenced by the detection of the *eaeA* gene, but neither *stx* gene fell within group A. Two isolates from one sample were classified as O157:H7, and based on PFGE, these isolates were of the same strain. Although virulence genes and *E. coli* O157:H7 were found in the manure composite and fecal samples, no virulent strains were found in any of the milk and milk filter samples.

Although *E. coli* mastitis does occur, it is generally believed that the majority of *E. coli* in bulk milk comes from fecal and

TABLE 7. ASSOCIATION OF VIRULENCE GENES WITH PHYLOGENETIC GROUPING IN DAIRY *ESCHERICHIA COLI* ISOLATES

Culture collection number	Source type	Sampling date	Sample source	Pulsed-field gel electrophoresis group	Virulence genes	Phylogenetic grouping
1001	Composite	2005	Postweaned calves	14	<i>stx2</i> <sup>+</sup> , <i>eaeA</i> <sup>+</sup> , <i>tir</i> <sup>+</sup>	D
1006	Composite	2005	Postweaned calves	14	<i>stx2</i> <sup>+</sup> , <i>eaeA</i> <sup>+</sup> , <i>tir</i> <sup>+</sup>	D
1011	Composite	2005	Postweaned calves	79	<i>eaeA</i> <sup>+</sup>	A
1014	Composite	2005	Postweaned calves	79	<i>eaeA</i> <sup>+</sup>	A
1553	Fecal	2006	Cow 525	30	<i>stx2</i> <sup>+</sup>	B1
1575	Fecal	2006	Cow 524	57	<i>stx2</i> <sup>+</sup>	B1
1585	Fecal	2006	Cow 524	93	<i>stx1</i> <sup>+</sup> , <i>stx2</i> <sup>+</sup>	B1
1590	Fecal	2006	Cow 524	57	<i>stx2</i> <sup>+</sup>	B1
1736	Fecal	2006	Cow 446	75	<i>stx1</i> <sup>+</sup>	D

environmental contamination introduced during the milking process. This particular farm has consistently demonstrated good hygienic practices (Van Kessel *et al.*, 2008), and the milk coliform counts are consistently very low. While overall the majority of isolates from the milk were unique to milk, some overlap of strains was observed with isolates from feces, water, and manure composites. Similar results were seen for the milk filter, but there was no overlap between the milk and milk filter isolates. Presumably, a more complete description of the *E. coli* population in the feces and environment would have resulted in more overlap between these compartments and the milk.

### Conclusion

In this study the level of diversity observed within the *E. coli* population on a single farm suggests that assigning a single environmental isolate to a particular farming operation would require the testing of an impractical number of isolates. Animal agriculture is frequently held responsible for contamination of potable and recreational waters with zoonotic bacterial pathogens that represent a major public health concern. Coliforms, particularly *E. coli*, have been used as indicators of fecal contamination of water, and source tracking using these indicator organisms and molecular techniques such as PFGE, ribotyping, and others has been proposed as a way of assigning responsibility for bacterial contamination of the environment. The data presented here suggest that it would be extremely difficult, in the absence of a unique isolatable factor, to track an individual strain back to its source. Success in the isolation and tracking of *E. coli* O157:H7 in clinical and environmental samples has been in part due to unique phenotypic characteristics (cefixime and tellurite resistance, inability to ferment sorbitol) that can be exploited in the isolation process. The overall environment of the modern dairy farm is complex. Despite the fact that animals in a free stall operation as described here are compartmentalized in groups based on production and nutritional needs, there is movement of animals between groups and movement of the groups to different areas of the barn during milking or while the barns are being cleaned. It is also an open system with bacterial inputs from humans, purchased animals, wildlife, and so forth. Understanding the relationship between all these inputs and the bacteria that ultimately contaminate the milk in the bulk tank or the environment is extremely challenging.

### Acknowledgments

The authors gratefully acknowledge the producer for providing us with access to his herd and the assistance of our Regional Dairy Quality Management Alliance (RDQMA) cooperators at The Pennsylvania State University. We also acknowledge the technical assistance of Thomas Jacobs, Jr., Claudia Lam, Crystal Rice-Trujillo, Jackie Sonnier, and Sam Shen.

### Disclaimer

Mention of a trade name, proprietary product, or specific equipment does not constitute a guarantee or warranty by the U.S. Department of Agriculture, and does not imply its approval to the exclusion of other products that may be suitable.

### Disclosure Statement

No competing financial interests exist.

### References

- Arthur TM, Bosilevac JM, Nou X, Shackelford SD, Wheeler TL, and Koohmaraie M. Comparison of the molecular genotypes of *Escherichia coli* O157:H7 from the hides of beef cattle in different regions of North America. *J Food Prot* 2007;70:1622–1626.
- Avery SM, Liebana E, Hutchison ML, and Buncic S. Pulsed field gel electrophoresis of related *Escherichia coli* O157 isolates associated with beef cattle and comparison with unrelated isolates from animals, meats and humans. *Int J Food Microbiol* 2004;92:161–169.
- Clermont OS, Bonacorsi S, and Bingen E. Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* 2000;66:4555–4558.
- Duriez P, Clermont O, Bonacorsi S, Bingen E, Chaventre A, Elion J, Picard B, and Denamur E. Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology* 2001;147:1671–1676.
- EPA Office of Science and Technology. Improved Enumeration Methods for the Recreational Water Quality Indicators: Enterococci and *Escherichia coli*. EPA/821/R-97/004, Washington, DC, 2000. Available at: [www.epa.gov/nerlcwww/RecManv.pdf](http://www.epa.gov/nerlcwww/RecManv.pdf).
- Escobar-Paramo P, Clermont O, Blanc-Potard AB, Bui H, Le BC, and Denamur E. A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. *Mol Biol Evol* 2004;21:1085–1094.
- Faith NG, Shere JA, Brosch R, Arnold KW, Ansay SE, Lee MS, Luchansky JB, and Kaspar CW. Prevalence and clonal nature of *Escherichia coli* O157:H7 on dairy farms in Wisconsin. *Appl Environ Microbiol* 1996;62:1519–1525.
- Girardeau JP, Dalmasso A, Bertin Y, Ducrot C, Bord S, Livrelli V, Vernozy-Rozand C, and Martin C. Association of virulence genotype with phylogenetic background in comparison to different seropathotypes of shiga toxin-producing *Escherichia coli* isolates. *J Clin Microbiol* 2005;43:6098–6107.
- Hancock DD, Besser TE, and Rice DH. Ecology of *Escherichia coli* O157:H7 in cattle and impact of management practices. In: *Escherichia coli O157:H7 and Other Shiga Toxin-Producing E. coli Strains*. Kaper JB and O'Brien AD (eds.). Washington, DC: ASM Press, 1998, pp. 85–91.
- Herzer PJ, Inouye S, Inouye M, and Whittam TS. Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *J Bacteriol* 1990;172:6175–6181.
- Higgins J, Hohn C, Hornor S, Frana M, Denver M, and Joerger R. Genotyping of *Escherichia coli* from environmental and animal samples. *J Microbiol Methods* 2007;70:227–235.
- Houser BA, Donaldson SC, Padte R, Sawant AA, DebRoy C, and Jayarao BM. Assessment of phenotypic and genotypic diversity of *Escherichia coli* shed by healthy lactating dairy cattle. *Foodborne Pathog Dis* 2008;5:41–51.
- Hunter PR and Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol* 1988;26:2465–2466.
- Jarvis GN, Kizoulis MG, Diez-Gonzalez F, and Russell JB. The genetic diversity of predominant *Escherichia coli* strains isolated from cattle fed various amounts of hay and grain. *FEMS Microbiol Ecol* 2000;32:225–233.

- Karns JS, Van Kessel JS, McClusky BJ, and Perdue ML. Incidence of *Escherichia coli* O157:H7 and *E. coli* virulence factors in US bulk tank milk as determined by polymerase chain reaction. *J Dairy Sci* 2007;90:3212–3219.
- Lahti E, Ruoho O, Rantala L, Hannonen ML, and Honkanen-Buzalski T. Longitudinal study of *Escherichia coli* O157 in a cattle finishing unit. *Appl Environ Microbiol* 2003;69:554–561.
- Leung KT, Mackereth R, Tien YC, and Topp E. A comparison of AFLP and ERIC-PCR analysis for discriminating *Escherichia coli* from cattle, pig and human sources. *FEMS Microbiol Ecol* 2004;47:111–119.
- Lukinmaa S, Nakari UM, Eklund M, and Siitonen A. Application of molecular genetic methods in diagnostics and epidemiology of food-borne bacterial pathogens. *APMIS* 2004;112:908–929.
- McDaniels AE, Rice EW, Reyes AL, Johnson CH, Haugland RA, and Stelma GN Jr. Confirmational identification of *Escherichia coli*, a comparison of genotypic and phenotypic assays for glutamate decarboxylase and beta-D-glucuronidase. *Appl Environ Microbiol* 1996;62:3350–3354.
- Ochman H and Selander RK. Standard reference strains of *Escherichia coli* from natural populations. *J Bacteriol* 1984;157:690–693.
- Ribot EM, Fair MA, Gautom R, Cameron DN, Hunter SB, Swaminathan B, and Barrett TJ. Standardization of pulsed-field gel electrophoresis protocols for the subtyping of *Escherichia coli* O157:H7, *Salmonella*, and *Shigella* for PulseNet. *Foodborne Pathog Dis* 2006;3:59–67.
- Schleif RF and Wensink PC. *Practical Methods in Molecular Biology*. New York: Springer-Verlag, 1981.
- Shere JA, Bartlett KJ, and Kaspar CW. Longitudinal study of *Escherichia coli* O157:H7 dissemination on four dairy farms in Wisconsin. *Appl Environ Microbiol* 1998;64:1390–1399.
- Thorpe CM, Ritchie JM, and Acheson DWK. Enterohemorrhagic and other shiga toxin-producing *Escherichia coli*. In: *Escherichia coli Virulence Mechanisms of a Versatile Pathogen*. Donnenberg MS (ed.). New York: Academic Press, 2002, pp. 119–154.
- Van Kessel JS, Karns JS, Wolfgang DR, Hovingh E, and Schukken YH. Longitudinal study of a clonal, subclinical outbreak of *Salmonella enterica* subsp. *enterica* serovar Cerro in a U.S. dairy herd. *Foodborne Pathog Dis* 2007;4:449–461.
- Van Kessel JS, Karns JS, Wolfgang DR, Hovingh E, Jayarao BM, Van Tassell CP, and Schukken YH. Environmental sampling to predict fecal prevalence of *Salmonella* in an intensively monitored dairy herd. *J Food Prot* 2008;71:1967–1973.
- Welinder-Olsson C, Badenfors M, Cheasty T, Kjellin E, and Kaijser B. Genetic profiling of enterohemorrhagic *Escherichia coli* strains in relation to clonality and clinical signs of infection. *J Clin Microbiol* 2002;40:959–964.
- Wells JG, Shipman LD, Greene KD, Sowers EG, Green JH, Cameron DN, Downes FP, Martin ML, Griffin PM, and Ostroff SM. Isolation of *Escherichia coli* serotype O157:H7 and other Shiga-like-toxin-producing *E. coli* from dairy cattle. *J Clin Microbiol* 1991;29:985–989.
- Whipp SC, Rasmussen MA, and Cray WC Jr. Animals as a source of *Escherichia coli* pathogenic for human beings. *J Am Vet Med Assoc* 1994;204:1168–1175.
- Wilkerson VA, Mertens DR, and Casper DP. Prediction of excretion of manure and nitrogen by Holstein dairy cattle. *J Dairy Sci* 1997;80:3193–3204.
- Zadoks RN and Schukken YH. Use of molecular epidemiology in veterinary practice. *Vet Clin. N Am Food Anim Pract* 2006;22:229–261.

Address correspondence to:

Jo Ann S. Van Kessel, Ph.D.

Environmental Microbial Safety Laboratory

Animal and Natural Resources Institute

Agricultural Research Service

U.S. Department of Agriculture

Building 173, BARC-East

Beltsville, MD 20705

E-mail: joann.vankessel@ars.usda.gov

