Incidence of *Escherichia coli* O157:H7 and *E. coli* Virulence Factors in US Bulk Tank Milk as Determined by Polymerase Chain Reaction

J. S. Karns,* J. S. Van Kessel,* B. J. McClusky,† and M. L. Perdue*3

*Environmental Microbial Safety Laboratory, USDA-ARS, Beltsville, MD 20705
†USDA Animal and Plant Health Inspection Service-Veterinary Services, Center for Epidemiology and Animal Health, Fort Collins, CO 80526

**ABSTRACT**

Samples of bulk tank milk from dairies across the United States, taken as part of the National Animal Health Monitoring Dairy 2002 survey, were analyzed for the presence of several genes encoding virulence factors associated with enterohemorrhagic forms of *Escherichia coli* (EHEC) using real-time and conventional PCR assays. Samples from 859 farms in 21 states were collected and enriched in EC medium at 42.5 °C to amplify any *E. coli* present, and DNA was isolated from the resulting biomass. The *eaeA* gene encoding intimin, a virulence factor associated with enteropathogenic forms of *E. coli* and EHEC, was detected in 199 (23%) of the samples. Thirty-six samples (4.2%) were positive for *eaeA*, the gamma allele of the translocated intimin receptor (γ-tir), found in EHEC strains of O157:H7, and one or both shiga-like toxin genes (*stx1* and *stx2*), a combination that may be indicative of the presence of O157:H7 EHEC. Testing these 36 samples with a commercially available real-time PCR kit for detection of O157:H7 indicated that 5 samples could be contaminated with O157:H7. A multiplex PCR to detect the presence of *fliC*, *rfbE*, and *hlyA* genes found in O157:H7 reduced to 2 (0.2% of all samples) the number of samples likely to be contaminated with this organism. A strain of O157:H7 (*eaeA*+, γ-tir+, *stx2*+, *rfbE*+, *fliC*+, *hlyA*+) was subsequently isolated from one sample. Thirty-four *eaeA*-positive samples did not contain detectable γ-tir but did contain one or both of the *stx* genes suggesting the presence of EHEC strains other than O157:H7. These results indicate a low incidence of O157:H7 in bulk tank milk but suggest that a risk from other enteropathogenic and EHEC forms of *E. coli* may exist and that PCR targeting virulence factors associated with highly pathogenic forms of *E. coli* may be an effective means of detecting potential dangers in raw milk.

**Key words:** enterohemorrhagic *Escherichia coli*, foodborne disease, polymerase chain reaction

**INTRODUCTION**

Prevention of foodborne illness is a global concern that affects all aspects of food production, including dairy. Dairy cattle are known reservoirs for zoonotic bacteria and milk and milk products have been implicated in outbreaks of foodborne illness (Wells et al., 1991; Zhao et al., 1995; Wallace, 1999). These outbreaks are generally caused by improperly pasteurized products, postpasteurization contamination, or by raw milk and raw milk products. *Salmonella* is probably the most common human pathogen associated with dairy products but other bacteria such as *Listeria monocytogenes*, *Yersinia* spp., *Campylobacter* spp., and pathogenic *Escherichia coli* are also present in dairy production environments and have been associated with milk and milk products (Ryser, 2001; Woteki and Kineman, 2003). Although milk can be contaminated with bacterial pathogens via shedding in the udder, these zoonotic pathogens are not commonly associated with mastitis. Fecal contamination is considered to be the most common route of bulk tank milk contamination with these organisms.

*Escherichia coli* live commensally in the gastrointestinal tract of most mammals, including humans, without causing disease. However, a small fraction of *E. coli* are human pathogens and have been implicated in foodborne illness with increasing frequency over the last 2 decades. Ingestion of enteropathogenic *E. coli* (EPEC) can result in a mild gastrointestinal disease that is relatively self-limiting but a subset of the EPEC, enterohemorrhagic *E. coli* (EHEC), can cause hemolytic-uremic syndrome (HUS), a serious, potentially fatal illness (O’Brien and Kaper, 1998). The most common human clinical EHEC isolated in the United States is
O157:H7 (Banatvala et al., 2001; Woteki and Kineman, 2003). Ground beef, a major by-product of the dairy industry, has been frequently associated with outbreaks of E. coli O157:H7 and milk-related outbreaks have also been reported, although these are less common (USDA:APHIS, 1997; Woteki and Kineman, 2003). The milk outbreaks were due to consumption of raw milk or milk that was contaminated postpasteurization (Ryser, 2001).

Escherichia coli O157:H7 has been frequently identified in cattle feces but, for the most part, adult cattle are asymptomatic shedders of this organism (Dean-Nystrom et al., 1999). As a result, identification of infected animals is difficult and there are no known mitigation methods to reduce or eliminate this organism from production environments. Therefore, it may be important from a public health standpoint to determine the prevalence of, and perhaps monitor the presence of, EHEC in the US raw milk supply. A few regional studies indicate that the prevalence of EHEC in bulk tank milk is relatively low. Shiga toxin-producing E. coli were detected in 3.8% of bulk milk samples from South Dakota and Minnesota (Jayarao and Henning, 2001) and Steele et al. (1997) reported that 0.87% of bulk tank samples collected in Ontario contained verotoxigenic E. coli. The E. coli O157:H7 serotype was detected in 0.75% (2/268) of raw milk samples obtained from dairy farms in eastern Tennessee (Murinda et al., 2002).

Isolation of E. coli O157:H7 from a biological sample is often difficult due to the large background of non-pathogenic E. coli. Standard plating techniques that take advantage of differences in sorbitol fermentation or antibiotic resistance still yield numerous non-O157:H7 isolates. Additionally, O157:H7 is only one of several potential EHEC serotypes, so methods that are designed to look solely for this serotype may overlook other highly pathogenic serotypes. To our knowledge, there is no single molecular or immunological method that can detect low levels of EHEC in a sample, such as milk, with a diverse microbial population. Therefore, we have analyzed samples from a national survey of bulk tank milk for the presence of several virulence genes associated with EHEC using PCR with the objective of determining the incidence of EHEC O157:H7 in US raw milk and to establish an idea of the levels of contamination by other EHEC serotypes.

**MATERIALS AND METHODS**

The milk samples used in this study were collected during the National Animal Health Monitoring System (NAHMS) Dairy 2002 survey. The states, the breakdown into regions for this study, and the number of samples from each region, are shown in Table 1. The methods for collection, shipping, and analysis of bulk tank milk samples have been described previously (Van Kessel et al., 2004). For enrichment of E. coli, 5 to 10 mL of milk was added to 95 mL of EC medium (Difco, Detroit, MI). The variation in volume was due to variation in available sample. Enrichment bottles were incubated at 42.5°C for 14 to 18 h. Two milliliters of the enrichment was centrifuged (13,000 × g, 2 min), the supernatant removed, and the resulting pelletted material was suspended in 0.5 mL of a 1:1 mixture of the 2× freezing medium described by Schleif and Wensink (1981) and Lennox broth (Gibco Laboratories, Long Island, NY), and frozen at –80°C. To reculture a sample, the frozen material was thawed and mixed; 0.1 mL was removed and added to 5 mL of Lennox broth and the cells grown at 37°C overnight.

For PCR analysis, enriched samples (1.5 mL) were centrifuged (13,000 × g) in microcentrifuge tubes, the supernatants were discarded, and the pellets were stored frozen at –20°C. The DNA was extracted from the bacterial pellets using 200 μL of InstaGene Matrix (BioRad Laboratories, Hercules, CA) following the manufacturer’s directions. The DNA preparations were stored at –20°C and were analyzed for the presence or absence of virulence genes associated with pathogenic E. coli via real-time or conventional PCR at a later date.

Immunomagnetic bead separation (IMS) was used to attempt to extract O157:H7 from samples that were implicated by PCR analysis. One milliliter of recultured milk enrichment sample was centrifuged for 1 min at 13,000 × g and the pellet suspended in 1 mL of PBS-2 [150 mM potassium phosphate (pH 7.2) and 150 mM NaCl]. To this cell suspension was added 40 μL of streptavidin beads (2.8 μm; Dynalbeads M-280, Dynal A.S., Oslo, Norway) complexed to monoclonal anti-E. coli O157 antibodies (BioDesign International, Saco, ME).
Figure 1. Schematic of the tiered PCR scheme used to detect virulence factors associated with *Escherichia coli* O157:H7 and other pathogenic *E. coli* in bulk tank milk.

The mix was incubated for 1 h at room temperature with shaking. This incubation mix contained approximately $10^7$ beads. Beads were condensed using a magnetized holder, suspended in 1 mL of PBS-2, and shaken for 15 min to dislodge poorly bound cells, extracted again, and resuspended in 1 mL of PBS-2. Then, 100 μL was plated onto sorbitol-MacConkey agar plates containing cefixime and tellurite by using an Autoplate 4000 spiral plater (Spiral Biotech, Norwood, MA). Plates were incubated overnight at 37°C. Sorbitol-nonfermenting colo-
nies were further analyzed by plating on MacConkey, sorbitol-MacConkey, and Simmons citrate agar. Colonies that fermented lactose, but not sorbitol or citrate were considered potential O157:H7 isolates. Agglutination tests on presumptive isolates of O157:H7 were performed using a RIM *E. coli* O157:H7 latex test kit (Remel, Lenexa, KS).

A tiered PCR assay was used to examine nucleic acids extracted from enrichments of each milk sample for the presence of virulence factors associated with pathogenic forms of *E. coli* (Figure 1). All 859 samples were assayed with a TaqMan real-time PCR assay designed to detect all alleles of the eaeA gene encoding intimin. The eaeA gene sequence was extracted from an *E. coli* O157:H7 genome fragment in GenBank (Accession no. AF071034) and is the reverse complement of bp 16482 to 19286. The extracted O157:H7 sequence was compared with eaeA sequences from other *E. coli* serotypes using CLUSTAL W (Thompson et al., 1994). The primer/probe set was selected for a conserved region of eaeA to detect a broad range of enterohemorrhagic and enteropathogenic *E. coli*. The primers JKP11 (GGCGATTACGCCGAAGATACC) and JKP12 (CCAGTGAACTACCGTCAAAGTTATTACC) are located at bp 1097 to 1115 and 1144 to 1124, respectively, relative to the first base of the ATG codon representing the site of initiation of translation of the gene. The TaqMan probe JKTM43 (CGCCTTACTGCGCGCTTGG) is located at bp 1047 to 1066. This assay was used as a PCR positive control to assure that there was no inhibition of PCR by substances in the nucleic acid extracts of enrichment cultures.

The eaeA, tir, and stx assays were performed using TaqMan Master Mix (Applied Biosystems, Foster City, CA) for 50 μL reactions in an ABI 7700 Sequence Detector (Applied Biosystems) or a Stratagene MX4000 (Stratagene, La Jolla, CA) with 300 nM of reverse primer and 250 nM of probe. The lacZ assays were performed in a similar fashion except that 900 nM of reverse primer (JKP42) was used. All PCR reactions also contained 50 μg/mL of BSA. The thermal program used on both instruments was 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. Fluorescence data were obtained during the annealing and extension steps. A sample was considered positive when the fluorescence reading crossed a threshold value at an amplification cycle (Ct) earlier than a no-template control. All primers and probes were obtained from Applied Biosystems. All probes were labeled with 6-FAM at the 5′-end and TAMRA at the 3′-end.

The O157:H7 detection kit from Idaho Technologies (Salt Lake City, UT) was used to further characterize samples that gave positive signals for eaeA and γ-tir using the Ruggedized Advanced Pathogen Identification Device (RAPID). Two microliters of sample was added to a 20-μL reaction. Additional PCR analysis was done on select samples using multiplex primer set C described by Wang et al. (2002). Sequencing of the DNA was done on an ABI 3100 Genetic Analyzer using the PCR primers and Big Dye sequencing kits (Applied Biosystems).

Chi-square goodness of fit tests were used to test the hypothesis that virulence factors in raw milk were uniformly distributed across geographical region, and among samples with various somatic cell counts or degrees of coliform contamination. The GENMOD procedure of SAS (SAS Institute, Cary, NC) was used to analyze binomial distributed data with a logit-link function. The statistical model included linear and quadratic regression terms for SCC or coliforms on the logit scale.

### RESULTS

Out of *E. coli* enrichments from 859 bulk tank milk samples taken from 21 states in the United States, 199 (23.2%) were positive for the presence of the intimin...
gene\textit{ (eaeA)\textit{, indicating the presence of possible pathogenic forms of \textit{E. coli}. Further examination of these 199 enrichments demonstrated that 36 (4.2\%) were also positive for the $\gamma$ allele of the translocated intimin receptor gene and one or both shiga-like toxin genes ($stx_1$, $stx_2$), indicating possible contamination with the EHEC O157:H7 strain of \textit{E. coli}. Analysis of raw milk samples seeded with O157:H7 indicated that the detection limit for the PCR reactions was between 0.1 and 1 cell/mL when single PCR reactions were run. The PCR of $\textit{lacZ}$ indicated there was no inhibition of amplification in any sample. Analysis of these 36 $\textit{eaeA}^+\textit{tir}^+\textit{stx}^+$ samples with a commercially available PCR assay for detection of O157:H7 reduced the number of samples likely to have O157:H7 contamination to 5 (0.6\%). Examination of these 5 samples by conventional multiplex PCR indicated that only 2 samples (0.02\%) had a high probability of O157:H7 contamination based on the presence of the rfb\textit{O157} (O157 serotype), $\textit{fliC}$ (H7 flagellar serotype), and $\textit{hlyA}$ (EHEC hemolysin) genes. Culture of these samples after IMS to concentrate any O157:H7 led to the isolation of a sorbitol-nonfermenting, cefixime- and telurite-resistant, $\textit{eaeA}^+\textit{tir}^+\textit{stx}^2\textit{rfb\textit{O157}}^+\textit{fliC}^+\textit{hlyA}^+$ strain of \textit{E. coli} from one sample. Agglutination assays confirmed that the strain was O157:H7.

It is important to note that PCR assays on EC enrichments detect the target genes in mixed cultures; thus, although the presence of O157:H7 would yield PCR reactions that were $\textit{eaeA}^+\textit{tir}^+\textit{stx}^+$, the presence of these 3 genes in a sample is only suggestive of the presence of the organism. Detection by PCR of this or other combinations of $\textit{eaeA}$ and $\textit{stx}$ in the EC-enriched biomass of a milk sample may be suggestive of the presence of EHEC forms of \textit{E. coli} other than O157:H7. The $\textit{eaeA}$ gene was detected in combination with one or both $\textit{stx}$ genes in 70 of the 859 bulk tank milk samples (8.1\%). Comparison of the incidence of the $\textit{eaeA}$, $\textit{stx}$, and $\gamma$-\textit{tir} genes between samples from different regions of the United States is shown in Figure 2. The incidence of $\textit{eaeA}$ was highest in samples from the Western and Northeast regions and lowest in samples from the Midwestern region. Chi-square goodness-of-fit tests ($\chi^2$ GOF) of uniform incidence indicated a lack of uniformity ($P = 0.025$) across regions for the incidence of $\textit{eaeA}$ in \textit{E. coli} enrichments from bulk tank milk. The Midwest region also yielded a lower percentage of enrichment cultures with a combination of $\textit{eaeA}$ and $\textit{stx}$ than did the other 3 regions. In general, one or both of the $\textit{stx}$ genes was found in less than half of the samples that were $\textit{eaeA}$ positive. The $\chi^2$ GOF tests indicated a uniform distribution ($P = 0.38$) of $\textit{stx}$ across production regions of enrichments containing the $\textit{eaeA}$ gene. The combination of $\textit{eaeA}$, $\gamma$-\textit{tir}, and $\textit{stx}$ was also lower in the Midwestern region; however, $\chi^2$ GOF tests indicated that the distribution of enrichments containing these 3 factors was uniform ($P = 0.58$) across the regions.

Figure 3 shows the relationship between the prevalence of the 3 virulence genes in enrichments with the total coliform count determined for the original milk samples. Generally, a greater incidence of $\textit{eaeA}$ was

![Figure 2. Prevalence of Escherichia coli virulence genes in DNA extracted from EC broth enrichments of bulk tank milk samples from NAHMS 2002 Dairy Survey by region. Shown is the incidence of all samples showing a positive PCR reaction for intimin (\textit{eaeA}) and the subset of those samples that also contain one or both of the shiga-toxin (\textit{stx}) genes. The subset of $\textit{eaeA}^+\textit{stx}^+$ samples that also contained the $\gamma$-allele of the translocated intimin receptor ($\gamma$-\textit{tir}) is also shown. Regions: \textit{W} = West; \textit{MW} = Midwest; \textit{NE} = Northeast; \textit{SE} = Southeast.](image1)

![Figure 3. Prevalence of Escherichia coli virulence genes in DNA extracted from EC broth enrichments of bulk tank milk samples from NAHMS 2002 Dairy Survey vs. the number of coliform bacteria detected in a sample.](image2)
seen in samples with higher coliform counts. No such trend was seen for combinations of eaeA and stx, or eaeA, stx, and γ-tir. The χ² GOF tests indicated a significant lack of uniformity in the incidence of eaeA⁺ enrichments (P < 0.005) across coliform count categories whereas the distribution of eaeA⁺stx⁻γ-tir⁺ enrichments was somewhat more uniform (P = 0.08). The distribution of eaeA⁺stx⁺ enrichments appeared to be uniform (P = 0.22) across coliform count categories. The χ² GOF results were further supported by logistic regression performed using the GENMOD module (SAS Institute), which indicated a highly significant linear relationship (P = 0.0017) between coliform count and the presence of eaeA.

The relationship between SCC and the prevalence of eaeA, stx, and γ-tir is shown in Figure 4. Samples with higher SCC generally were more likely to yield enrichment cultures in which the 3 virulence factors were detected. The χ² GOF tests indicated a significant lack of uniformity in the distribution of eaeA⁺ (P = 0.006) and eaeA⁺stx⁺ (P = 0.040) enrichments across the SCC categories. The distribution of eaeA⁺stx⁺γ-tir⁺ enrichments was somewhat more uniform (P = 0.077) over the SCC categories. Logistic regression confirmed a highly significant log-linear relationship (P = 0.0031) between SCC and the presence of eaeA in enrichments and a significant logit-linear relationship between SCC and the incidence of eaeA⁺stx⁻γ-tir⁺ enrichments (P = 0.033).

To obtain an idea of the occurrence of the genes used for detection of E. coli O157:H7 in other milk samples, the multiplex PCR described by Wang et al. (2002) was run on 15 samples that did not seem to be contaminated with O157:H7. Nine samples were eaeA⁺γ-tir⁺stx⁺ but were negative when tested with the Idaho Technologies O157:H7 kit; 4 samples were eaeA⁺γ-tir⁻stx⁺ and were negative when tested with the Idaho Technologies O157:H7 kit; and 2 samples were eaeA⁺γ-tir⁻stx⁻, with 1 showing a positive and 1 a negative result with the Idaho Technologies kit. The hlyA gene, indicating the presence of E. coli strains containing the EHEC plasmid, was found in 5 of these enrichments, all of which were stx⁺. The fliC gene encoding the flagellar protein responsible for the H7 serotype was found in 9 samples, one of which was stx⁻. None of these samples contained the rfbO157 gene. One sample (eaeA⁺γ-tir⁺) was rfbO157⁺ and fliC− suggesting possible contamination by an O157:H7 strain but was stx⁻ and hlyA⁺, suggesting an absence of any EHEC forms. The DNA sequences of the bands corresponding to the PCR products of rfbO157, fliC, and hlyA extracted from agarose gels were identical (98% or greater) to sequences of the corresponding O157:H7 genes.

**DISCUSSION**

The routine identification of pathogenic forms of E. coli in agricultural samples is complex and time consuming. Although methods for isolating EHEC O157:H7 exist, the isolation of the organism is generally more difficult with agricultural samples because the diversity of microorganisms often results in cross reactions. Identification of other EHEC serotypes is even more difficult because they lack the metabolic defect and antibiotic resistances found in O157:H7 strains. The use of PCR for the detection of virulence genes associated with pathogenicity in E. coli offers one way of determining the likelihood of contamination of a sample by pathogenic forms of E. coli regardless of serotype. In general, EHEC strains contain the locus of enterocyte effacement, of which eaeA is a part, and one or both shiga-like toxin genes. Thus, the detection of eaeA together with stx1 or stx2 in a milk sample could be considered suggestive of contamination by a pathogenic strain of E. coli. Because the survey described here was done on enrichments of the general coliform population in each sample, the PCR assays only indicate the presence of each gene within this population and do not discern whether the genes are present in a single, highly pathogenic form of E. coli. However, detection of these genes within this population should allow some assessment of the risks caused by these organisms in agricultural products.

Based upon this study, the incidence of EHEC O157:H7 in US bulk tank milk is very low. Only 2 of 859 samples (0.02%) were shown to be contaminated by this organism by PCR and a culture was only ob-
tained from 1 sample. However, the possibility of contamination by other EHEC serotypes is somewhat greater. The PCR assays on a subset of samples that did not appear to be contaminated with O157:H7 but were shown to contain both eaeA and stx genes showed that 5 of 13 (38.5%) also contained hlyA, indicating potential for EHEC in those samples. Seventy-one of the 859 samples surveyed contained both eaeA and stx, and if 38.5% of them were hlyA positive the potential exists for 27 samples to contain EHEC. The perfect set of PCR assays for assessment of EHEC risks associated with raw milk remains to be determined; in this study several confirmatory assays were required to detect samples that contained O157:H7. Based upon this study, initial assay for the presence of eaeA and stx genes followed by PCR detection of hlyA in eaeA*stx* samples might be sufficient to suggest that a sample of raw milk may be contaminated with EHEC. However, there have been reports of an HUS-causing E. coli strain that lacks eaeA (Paton et al., 1999); although such a strain would be missed by the assays described herein, no such strain has been reported in cattle or in dairy products.

It is tempting to view the apparent relationships between SCC or coliform counts with the incidence of eaeA or eaeA-stx combinations in enrichments as signifying the potential for risk assessment through these easily measured parameters. Likewise, it is tempting to conclude that milk from certain regions may be less likely to contain pathogenic E. coli. However, of the 2 samples that were shown to contain O157:H7, one had a low coliform count of 16 cfu/mL but a high SCC of 761,000 cells/mL, whereas the other had a coliform count of 4,500 cfu/mL and a moderate SCC of 404,000 cells/mL. Additionally, one of these samples came from a farm in the Midwest, which had the lowest incidence of eaeA and eaeA stx containing samples. Thus, to our knowledge, no easily measured parameter adequately describes the risk of EHEC in milk.

Although contamination of dairy products currently accounts for a small percentage of foodborne illness in the United States, it is clear that raw milk consumption and the consumption of products made with raw milk presents some risk. Although proper pasteurization minimizes these risks to the public, there is a small but growing group of people that consume unpasteurized milk or milk products, either for practical (farm families) or cultural (soft ethnic cheeses) reasons, or because of perceived health benefits (Cody et al., 1999; Villar et al., 1999; Anonymous, 2002). This has lead to outbreaks of O157:H7 diarrheal disease and HUS due to consumption of contaminated raw milk (Keene et al., 1997).

Although the incidence of E. coli O157:H7 in the milk samples tested here seemed to be very low (0.02%), the populations of E. coli in a higher fraction of samples (8.3%) contained a combination of virulence genes suggestive of other EHEC forms. In addition, 15% of the milk samples were contaminated by E. coli strains that may be milder pathogenic forms because they contain the eaeA gene. Because we only assayed eaeA-containing samples for the presence of the genes encoding stx1 and stx2, we do not know the incidence of these toxin genes across all milk samples. The incidence of potentially pathogenic E. coli forms reported here, combined with the severity of the disease they can cause and the potential for this organism to grow in improperly stored raw milk and in products made from raw milk, means E. coli in raw milk represents a significant public health risk, particularly to more susceptible members of the population. There is evidence suggesting that the shedding of O157:H7 by cattle increases during summer months (Hancock et al., 1997; Van Donkersgoed et al., 1999; Dunn et al., 2004). The samples for this study were collected between March and June and therefore may underestimate the true prevalence of EHEC in bulk tank milk. Continuing surveys of milk will help estimate the true level of risk associated with these practices and may help to identify dairy management practices that minimize the contamination of bulk tank milk with zoonotic foodborne pathogens.

ACKNOWLEDGMENTS

We thank Thomas Jacobs Jr., Sumit Dua, and Claudia Lam for technical assistance.

REFERENCES

PATHOGENIC E. COLI IN BULK TANK MILK


