IMPROVING MOLECULAR DETECTION AND DISCRIMINATION OF *SALMONELLA* AND SHIGA TOXIN-PRODUCING *ESCHERICHIA COLI*

By

CHRIS TIMMONS

Bachelor of Science in Biology

University of Texas at Tyler

Tyler, Texas

2010

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2012

IMPROVING MOLECULAR DETECTION AND DISCRIMINATION OF *SALMONELLA* AND SHIGA TOXIN-PRODUCING *ESCHERICHIA COLI*

Thesis Approved:

Li Maria Ma, Ph.D.

Thesis Advisor

Jacqueline Fletcher, Ph.D.

Robert Allen, Ph.D.

Name: CHRIS TIMMONS

Date of Degree: DECEMBER, 2012

Title of Study: IMPROVING MOLECULAR DETECTION AND DISCRIMINATION OF *SALMONELLA* AND SHIGA TOXIN-PRODUCING *ESCHERICHIA COLI*

Major Field: ENTOMOLOGY AND PLANT PATHOLOGY

ABSTRACT: Foodborne illnesses caused by non-O157 Shiga toxin-producing *Escherichia coli* (STEC) and *Salmonella enterica* (*SE*) are significant health concerns and economic burdens worldwide. While *SE* is an established pathogen, causing more illnesses, hospitalizations, and deaths in the United States than any other foodborne bacterial pathogen, non-O157 STEC are emerging pathogens of growing concern. Multistate and multinational outbreaks of foodborne illness are frequently associated with *SE* and non-O157 STEC contaminated food. Effective surveillance of foodborne illnesses and investigations of foodborne illness outbreaks rely on rapid, robust, and sensitive methods for pathogen detection and strain discrimination. The objectives of this research were 1) to evaluate the effect on PCR sensitivity of adding a short, AT-rich overhanging nucleotide sequence (flap) to the 5' end of PCR primers specific for the detection of *Salmonella* and *E. coli* O157:H7 and 2) to develop a multiple-locus variable-number tandem repeat (VNTR) analysis (MLVA) method for strain discrimination of 6 major serogroups of non-O157 STEC. When targeting individual pathogens, end-point PCR assays using flap-amended primers were more efficient than non-amended primers, with 20.4% and 23.5% increases in amplicon yield for *Salmonella* and *E. coli* O157:H7, respectively. In multiplex PCR assays, a 10- to 100-fold increase in detection sensitivity was observed when the primer flap sequence was incorporated. The MLVA method developed for non-O157 STEC used 12 VNTR loci that were amplified in 3 multiplex PCR reactions and sized by multicolor capillary electrophoresis. All serogroups were differentiable by the method, as 4 of the 6 serogroups were clustered separately in a minimum spanning tree. The developed MLVA method was more discriminatory for serogroups O26, O111, O103, and O121 than it was for O45 and O145. Compared to pulsed-field gel electrophoresis (PFGE), the "gold standard" bacterial pathogen subtyping technique, the MLVA method exhibited higher discriminatory power for serogroup O26, a similar level of discrimination for serogroups O111, O103, and O121, and a lower level of discrimination for serogroups O45 and O145. The methods developed in this research have potential applications in the food and agricultural industries, foodborne outbreak surveillance and investigations, and biosecurity and microbial forensics.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

CHAPTER III

CHAPTER IV

LIST OF FIGURES

Figure Page

CHAPTER III

CHAPTER IV

Figure Page

CHAPTER I

INTRODUCTION

Foodborne illnesses caused by Shiga toxin-producing *Escherichia coli* (STEC) and *Salmonella enterica* (*SE*) are significant health concerns and economic burdens worldwide (Voetsch *et al.,* 2004; Lynch *et al.,* 2009; Mathusa *et al.* 2010; Newell *et al.,* 2010). In the United States alone, approximately 1.2 million illnesses, 43,000 hospitalizations, and 400 deaths occur each year as a result of STEC and nontyphoidal *Salmonella* infections (Scallan *et al.,* 2011). The fecal-oral route is the primary mode of transmission for these bacterial pathogens among humans and illness is usually a result of consumption of contaminated food or water (Tauxe, 2002; Viazis and Diez-Gonzalez, 2011; Karmali, 1989). While meat or animal products are most often implicated, in recent years an increasing incidence of STEC and *SE* infections have been linked to consumption of fresh produce and other plant products. Contamination of food products by human pathogens can occur at any point throughout the field production, harvesting, processing, transport, storage, and preparation steps involved in the increasingly global food supply. Multistate and multinational outbreaks of foodborne illness as a result of such contamination are not uncommon (Scallan *et al.,* 2011). Positive identification of the point of contamination through outbreak investigations by epidemiologists and public

health officials is often difficult and the tools necessary for timely correlation of clinical, environmental, and food isolates are not available for all foodborne pathogens, especially for emerging pathogens.

Effective surveillance of foodborne pathogens and investigations of foodborne illness outbreaks rely on rapid, robust, and sensitive methods for pathogen detection and strain discrimination. Traditional culture techniques and molecular-based methods, such as polymerase chain reaction (PCR), are commonly used for detecting foodborne pathogens in clinical, environmental, and food samples. Due to their speed and sensitivity, PCR-based methods provide distinct advantages in time sensitive outbreak investigations and are continually being improved to allow fewer target pathogens to be detected in less time (Levin, 2009). Small modifications in PCR methods can sometimes lead to significant improvements in foodborne pathogen detection speed and sensitivity.

In addition to detection, discrimination of highly clonal pathogenic bacteria at the strain level is critical for epidemiological investigations. Pulsed-field gel electrophoresis (PFGE) is the current "gold standard" bacterial pathogen strain discrimination technique. Although the strain discrimination capability of PFGE is well documented, the technique has several drawbacks. PFGE is a time-consuming and laborious method requiring a high level of technical skill and rigorous standardization to allow inter-laboratory data sharing. At times, PFGE also fails to provide optimal discrimination among closely related bacterial isolates (Hyytia-Trees *et al.,* 2006). PCR-based discrimination methods allow higher sensitivity, increased assay speed, simplified interpretation of results, and increased discriminatory power than PFGE.

Multiple locus variable-number tandem-repeat (VNTR) analysis (MLVA) is a PCR-based strain discrimination technique now being used by the Centers for Disease Control and Prevention (CDC) to augment PFGE data. MLVA often allows a higher level of strain discrimination than is possible by PFGE alone (Hyytia-Trees *et al.,* 2006). MLVA assays have been developed and validated for STEC O157:H7, *Salmonella* Typhimurium, and *Salmonella* Enteritidis. However, the numbers of illnesses caused by over 380 non-O157 STEC serotypes and over 1500 other *SE* serotypes have been increasing in the U.S. and worldwide and these serotypes have been involved in multiple disease outbreaks (World Health Organization, 1999; Mathusa *et al.,* 2010). In many countries, these serotypes cause the majority of STEC and *SE* infections (Johnson *et al.,* 2006; CDC, 2008). As of yet, no MLVA assay has been developed specifically for non-O157 serogroups or for most other significant *SE* serovars. Such methods would have been useful during a recent outbreak of foodborne illness originating in Germany involving a newly emerging STEC serogroup (Bezuidt *et al.,* 2011) and a recent outbreak of illness attributed to *Salmonella* Montevideo in the U.S. in which strain discrimination by PFGE was difficult (Lienau *et al.,* 2011).

The objectives of this research were 1) to improve the sensitivity of PCR for detection of *E. coli* O157:H7 and *Salmonella* using AT-rich 5' flap sequences on PCR primers and 2) to develop a MLVA assay for discrimination of 6 major serogroups of non-O157 STEC. The developed methods allow sensitive multiplex detection of STEC O157 and *Salmonella* as well as inter- and intra-serogroup discrimination of non-O157 STEC. The developed assays will be valuable in outbreak surveillance, outbreak investigations, and agricultural biosecurity and microbial forensics applications.

LITERATURE CITED

Bezuidt, O., R. Pierneef, K. Mncube, G. Lima-Mendez, and O.N. Reva. (2011). Mainstreams of horizontal gene exchange in Enterobacteria: Consideration of the outbreak of enterohemorrhagic *E. coli* O104:H4 in Germany in 2011. *PLoS ONE.* 6(10):e25702.

Centers for Disease Control and Prevention. (2008). Salmonella Surveillance: Annual Summary, 2006. US Department of Health and Human Services, Atlanta, GA.

Hyytiä-Trees, E., S.C. Smole, P.A. Fields, B. Swaminathan, and E.M. Ribot. (2006). Second generation subtyping: A proposed PulseNet protocol for multiple-locus variable-number tandem repeat analysis of Shiga toxin–producing *Escherichia coli* O157 (STEC O157). *Foodborne Pathog. Dis.* 3:118-131

Johnson, L.E., C.M. Thorpe, and C.L. Sears. (2006). The emerging clinical importance of non-O157 Shiga toxin—producing *Escherichia coli. Clin. Infect. Dis.* 43: 1587-1595.

Karmali, M.A. (1989). Infection by Verocytotoxin-producing *Escherichia coli*. *Clin. Microbiol. Rev.* 2:15-38.

Levin, R. (2009). The use of molecular methods for detecting and discriminating *Salmonella* associated with foods. A review. *Food Biotech.* 23:313-367.

Lienau, E.K., E. Strain, C. Wang, J. Zheng, A.R. Ottesen, C.E. Keys, R. Stones. (2011). Identification of a salmonellosis outbreak by means of molecular sequencing. *New Engl. J. Med.* 364: 981-982.

Lynch, M.F., E.M. Blanton, S. Bulens, C. Polyak, J. Vojdani, J. Stevenson, and E.D. Mintz. (2009). Typhoid fever in the United States, 1999-2006. *JAMA: J. Am. Med. Assoc.* 302:859-865.

Mathusa, E.C., Y. Chen, E. Enache, and L. Hontz. (2010). Non-O157 Shiga toxin-producing *Escherichia coli* in foods. *J. Food Protect.* 73: 1721-1736.

Newell, D.G., M. Koopmans, L. Verhoef, E. Duizer, A. Aidara-Kane, H. Sprong, and H. Kruse. (2010). Food-borne diseases — The challenges of 20 years ago still persist while new ones continue to emerge. *Int. J.Food Microbiol.* 139:S3-S15.

Scallan, E., R.M. Hoekstra, F.J. Angulo, R.V. Tauxe, M.-A. Widdowson, S.L. Roy, J.L. Jones, and P.M. Griffin. (2011). Foodborne illness acquired in the United States—major pathogens. *Emerg. Infect. Dis.* 17:7-15

Tauxe, R.V. (2002). Emerging foodborne pathogens. *Int. J. Food Microbiol.* 78:31-41.

Viazis, S. and F. Diez-Gonzalez. (2011). Chapter one - Enterohemorrhagic Escherichia coli: The Twentieth Century's Emerging Foodborne Pathogen: A Review, pp. 1-50. *In* L. S. Donald (ed.), Advances in Agronomy. Academic Press.

Voetsch, A.C., T.J. Van Gilder, F.J. Angulo, M.M. Farley, S. Shallow, R. Marcus, P.R. Cieslak, V.C. Deneen, R.V. Tauxe, and f.t.E.I.P.F.W. Group. (2004). FoodNet estimate of the burden of illness caused by nontyphoidal *Salmonella* infections in the United States. *Clin. Infec. Dis.* 38:S127-S134.

World Health Organization. (1999). World Health Organization, 1999. Zoonotic non-O157 Shiga toxin-producing *Escherichia coli* (STEC). Report of a WHO Scientific Working Group Meeting. Berlin, Germany, 22–23 June, 1998, WHO, pp 1-30.

CHAPTER II

LITERATURE REVIEW

Foodborne illness

As defined by the World Health Organization (WHO), foodborne illnesses are diseases caused by infectious or toxic agents that enter the body through contaminated food (WHO, 2007). Foodborne illness can be caused by bacteria, parasites, viruses, or toxins and poses a significant health concern and economic burden in the U.S. and worldwide (Hird *et al.,* 2009). Approximately 9.4 million illnesses, 55,961 hospitalizations and 1,351 deaths occur each year in the U.S., associated with 31 major foodborne pathogens (Scallan *et al.,* 2011). Direct and indirect annual costs associated with foodborne illness in the U.S. have been estimated to approach 35 billion dollars (Buzby and Roberts, 2009; Hoffman *et al.,* 2012). Nontyphoidal *Salmonella enterica* (*SE*) and Shiga toxin-producing *E. coli* (STEC) are two of the most commonly implicated bacterial foodborne pathogens (Voetsch *et al.,* 2004; Rangel *et al.,* 2005), causing an estimated 1,026,561 and 175,905 illnesses, respectively, in the U.S. annually (Scallan *et al.,* 2011). The rates of morbidity and mortality caused by *SE* and STEC are very similar. While STEC are slightly more virulent, the occurrence of illnesses caused by *SE* is much higher. *SE* has a moderate domestic mortality rate (estimated at 0.5% in the U.S.) but the

large number of infections results in this pathogen annually causing approximately 28% of foodborne related deaths in the U.S. *SE* causes more hospitalizations and deaths in the U.S. than any other foodborne pathogen (Scallan *et al.,* 2011).

Salmonella. *Salmonella* are rod-shaped, Gram-negative, peritrichous, non-spore forming, facultative anaerobic bacteria in the family *Enterobacteriaceae*. They are fairly ubiquitous in nature but the primary habitat is the intestinal tract of a variety of birds, reptiles, and mammals. According to the current nomenclature system used by the CDC and WHO, the genus *Salmonella* is divided into two species: *Salmonella enterica* and *Salmonella bongori* (Reeves *et al.,* 1989; Popoff and Le Minor, 1997), with *S. enterica* containing six subspecies (Brenner, 1998). The subspecies are designated by the Roman numerals I, II, IIIa, IIIb, IV, and VI corresponding to subspecies *enterica, salamae, arizonae, diarizonae, houtenae*, and *indica*, respectively.

Strains of *Salmonella* are characterized into serotypes or serovars (serological variants) according to antisera agglutination profiles corresponding to somatic (O) and flagellar (H) antigens. The O antigen is a polysaccharide component of lipopolysaccharides and the H antigen is the filamentous portion of flagella, composed of the protein flagellin. Differences in flagellin and the O antigen sugar identities, linkages, and bonding profiles provide the variability that allows subtyping (CDC, 2008). *Salmonella* commonly exhibits phase variation—a random switching between two flagellar phenotypes. As a result, a single genetically identical serovar may express either of two different H antigens at any given time, designated as Phase 1 or Phase 2

(CDC, 2006). *Salmonella* serovars often have unique names correlating to the location where they were first isolated.

While over 2500 serovars of *Salmonella* have been identified, the majority of human infections are caused by strains of *Salmonella enterica* subsp. *enterica* (Popoff and Le Minor, 1997; 2001), a taxon that includes over 1500 serovars (Popoff *et al.,* 2004). Unless otherwise noted, the names of clinically important serovars usually refer to *Salmonella enterica* subsp. *enterica* (*SE*). For example, *Salmonella* Typhimurium is the simplified form of *Salmonella enterica* subsp. *enterica* serovar Typhimurium.

The primary clinical syndromes associated with *SE* infection are enteric fever (also known as typhoid and paratyphoid fever) and gastroenteritis (also known as salmonellosis) (Miller and Peages, 2000; Ohl and Miller, 2001). Enteric fever, caused by *Salmonella* Typhi and Paratyphi, is the most serious syndrome but has been largely eradicated in the developed world (Crump *et al.,* 2004). Salmonellosis, caused by nontyphoidal *SE*, such as *S*. Enteritidis and *S*. Typhimurium, is the most common syndrome, and is usually associated with cellular internalization of the bacterium in the intestines (Francis *et al.,* 1992, 1993) and subsequent massive neutrophil recruitment to the infected cell (McCormick *et al.,* 1995; Santos *et al.,* 2001). These events result in localized cell death and significant water flow into the intestinal lumen, causing diarrhea (Hersch *et al.,* 1999; Cookson and Brennan, 2001). Other symptoms include fever, abdominal cramping, vomiting, and nausea (Coburn *et al.,* 2007). Severe infections and death are usually associated with the very young, very old, and immune-compromised (Celum *et al.,* 1987). Nontyphoidal *SE* cause more illnesses, hospitalizations, and deaths

than any other bacterial pathogen and no significant decline in illness occurrence has been observed in recent years (Scallan *et al.,* 2011).

According to the most recent annual CDC summary on *Salmonella* surveillance, the 10 most common nontyphoidal *SE* serovars in the U.S. in recent years are Typhimurium, Enteritidis, Newport, Heidelberg, Javiana, I4,[5],12:i:-, Montevideo, Muenchen, Oranienburg, and Mississippi, in the order of highest to lowest isolation frequency (CDC, 2008). Serovars Typhimurium, Enteritidis, Newport, and Heidelberg have been the most common serotypes since 1995, and *SE* serovars Typhimurium and Enteritidis have remained the most frequently isolated serovars since 1997. Although the total numbers of *SE* isolates in the U.S. have increased since 1996, isolation of Typhimurium and Enteritidis has substantially decreased, possibly as a result of preventative measures aimed at food sources commonly associated with these serovars (Mason, 1994; Houge *et al.,* 1997). Since foodborne pathogen surveillance began in the U.S. in 1996, the occurrence of nontyphoidal *SE* infections has remained relatively constant, while illnesses caused by similar foodborne pathogens have decreased (Scallan *et al.,* 2011).

Shiga toxin-producing *E. coli* (STEC). *Escherichia coli* are also rod-shaped, Gram-negative, peritrichous, non-spore forming, facultative anaerobic bacteria in the family *Enterobacteriaceae*. Six species of *Escherichia* exist but only *E. coli* is of significant clinical importance and associated with human illness (Chaudhury *et al.,* 1999). As with *Salmonella,* strains of *E. coli* are characterized into serogroups or serotypes according to antisera agglutination profiles corresponding to somatic (O),

flagellar (H), and sometimes capsular (K) surface antigens (Edwards and Ewing, 1972). The *E. coli* serogroup designation refers only to O antigen agglutination profiles while the serotype designation refers to O, H, and sometimes K antigen agglutination profiles.

A majority of *E. coli* serogroups are commensal in humans and are important constituents of the natural gut micro flora. However, several virulent *E. coli* subtypes exist and exhibit one of at least 5 different modes of pathogenesis in humans (Todar, 2012; Nataro and Kaper, 1998). Based on the mode of pathogenesis, virulent *E. coli* subtypes are referred to as Enterohemorrhagic *E. coli* (EHEC), Enteroaggregative *E. coli* (EAEC), Enteroinvasive *E. coli* (EIEC), Enterotoxigenic *E. coli* (ETEC), or Enteropathogenic *E. coli* (EPEC) (CDC, 2012). The most severe *E. coli*-induced human illness is caused by Shiga toxin-producing *E. coli* (STEC), a designation that can include both EHEC and EAEC (Scheutz *et al.,* 2011). STEC strains are associated mainly with at least two potent verocytotoxins (Nataro and Kaper, 1998), also called Shiga-like toxins due to their similarity to the primary toxin produced by *Shigella*, a human-limited bacterial pathogen very similar to EIEC (Pupo *et al.,* 1997). The primary response to Shiga toxins in humans is thrombotic microangiopathy in the renal glomeruli and sometimes other organs (Karmali *et al.,* 2010). Over 380 serotypes of STEC have been identified (World Health Organization, 1999) and infections caused by STEC are usually associated with diarrhea, hemorrhagic colitis, and hemolytic uremic syndrome (HUS) (Karmali *et al.,* 1985; Karmali, 1989). Similar to *Salmonella* infections, severe cases of STEC are usually associated with the very young, very old, and immune-compromised (Griffin *et al.,* 1988, Karmali, 1989; Gould *et al.,* 2009).

Shiga toxin production and HUS were only relatively recently associated with *E. coli* O157:H7 (Karmali *et al.,* 1983). First identified and isolated from meat, this pathogen continues to cause significant morbidity and mortality worldwide and has been isolated from a variety of food sources (Rangel *et al.,* 2005). Costs associated with STEC O157:H7 illnesses alone have been estimated at 344 million dollars in the U.S. (Frenzen *et al.,* 2005). Among STEC serotypes, O157:H7 causes the highest number of sporadic and outbreak related cases (Griffin and Tauxe, 1991) but causes a minority in the total number of illnesses caused by all other STEC serotypes (Johnson *et al.,* 2006, Hadler *et al.*, 2011). In a recent population-based study using 10 years' worth of data on the differences between O157 STEC and non-O157 STEC in regard to incidence of infection, severity of infection, and genetic similarities, it was found that (1) total incidence of all STEC infections decreased from 2000 to 2009, (2) the leading STEC serotypes are consistent in occurrence, (3) serotype O157 is associated with the most severe disease, and (4) there is a substantial amount of diversity in epidemiology and risk factors among the non-O157 STECs (Hadler *et al.,* 2011). Another recent report (Gilliss *et al.*, 2011) showed that the occurrence of O157 STEC has decreased since active surveillance began in 1996, and that in 2010 the goal for low incidence was met.

Salmonella **and STEC surveillance**

In the early 1990's, a markedly increased incidence of foodborne illness was observed in the U.S. and led to national initiatives aimed at reducing the future health and economic burden of such illnesses (CDC, 1994). Foodborne illness surveillance systems were developed to meet this need and are defined by the CDC as systems that "collect

and analyze morbidity (rate of disease incidence), mortality (rate of diseases resulting in death), and other relevant data and facilitate the timely dissemination of results to appropriate decision makers" (Sosin, 2003). FoodNet and PulseNet, two highly successful foodborne disease surveillance programs developed by the CDC in 1996, are the primary surveillance systems used in the U.S. and have been modeled by several other countries (Swaminathan *et al.,* 2006). Although the CDC began surveillance of foodborne bacterial pathogens in 1962, it was not until 1996 that a more complete understanding of the epidemiology and the true burden of STEC, *SE*, and many other foodborne pathogen infections in the U.S. became possible (Olsen *et al.,* 2001).

FoodNet (the Foodborne Disease Active Surveillance Network) is an active national surveillance program targeting 7 bacterial and 2 parasitic pathogens that are known to cause foodborne illnesses, as identified by the CDC's Emerging Infections Program, the U.S. Department of Agriculture (USDA), the U.S. Food and Drug Administration (FDA), and several state health departments. Accurately estimating the burden of foodborne illness in the U.S., investigating the sources of infection in outbreaks and sporadic cases, and building public health infrastructure for dealing with emerging foodborne illnesses are the main objectives of this program (CDC, 1997; Angulo *et al*., 1998; Voetsch *et al.,* 2004). FoodNet personnel located at state health departments regularly contact several state and regional clinical laboratories to collect reports of infections diagnosed in residents of these areas. The FoodNet surveillance area includes about 15% of the U.S. population and the information gathered is used to assess the impact of food safety initiatives and the burden of foodborne illness (CDC, 2012)

PulseNet (the National Molecular Subtyping Network for Foodborne Disease Surveillance) is a relatively passive surveillance system established by the CDC in collaboration with several cooperating state health departments. Through molecular subtyping of pathogens, PulseNet helps monitor the national incidence of illnesses caused by foodborne pathogens. The PulseNet database is made up of pulsed-field gel electrophoresis (PFGE) profiles submitted by state and local public health laboratories and federal food regulatory laboratories that perform molecular surveillance of foodborne infections (Swaminathan *et al.,* 2006). In recent years several countries in Europe, Latin America, Africa, and the Asia Pacific region have replicated or linked to the PulseNet network, thus providing an international approach for identification and response to international outbreaks of foodborne illness (Gerner-Smidt *et al.,* 2006; Swaminathan *et al.,* 2006). By maintaining and monitoring the PulseNet database, the CDC develops baselines of illness occurrence for specific strains of pathogens (correlating to specific PFGE profiles). Unexpected increases in the occurrence of specific PFGE profiles help epidemiologists identify clusters of foodborne illnesses that may or may not be epidemiologically related. Further investigation is needed to determine if an outbreak of foodborne illness has occurred. Accurate identification of sporadic as well as outbreak related illnesses are essential for improving this empirical system over time.

Sporadic cases. Like illnesses caused by many other foodborne pathogens, those caused by *SE* and STEC are usually sporadic in nature with only a small percentage of cases being part of an outbreak. For example, between 1998 and 2002 a total of 164,044 *SE* infections were reported to the National *Salmonella* Surveillance System, of which

only 16,821 were outbreak associated (Lynch *et al*., 2006). According to these data, as well as those from previous years, more than 80% of *Salmonella* infections are not associated with outbreaks. Similarly, sporadic cases of *E. coli* O157:H7 infections have been estimated to cause as much as 90% of the total *E. coli* O157:H7 infections in some developed countries (Thomas *et al.,* 1996). Even though sporadic cases make up the majority of observed cases in the U.S., it is often difficult or impossible to find the specific source of infection. Sporadic cases mainly help establish a baseline of background illness occurrence for more rapid identification of illness outbreaks. It is primarily through the investigation of foodborne illness outbreaks that the sources of foodborne infection and contaminated foods can be identified. Identification of the source of contamination and/or the route of contamination can help public health officials and regulatory agencies implement appropriate policies that will reduce the likelihood of similar illnesses occurring in the future.

Foodborne illness outbreaks. An outbreak can be defined as the occurrence of two or more cases of foodborne illness that can be traced to a common source. In most developed countries, outbreaks are usually detected and monitored by surveillance systems. In the U.S., bacterial foodborne disease outbreaks are identified primarily by the PulseNet network. A statistically significant increase in the number of PFGE profiles added to the PulseNet database over a given time and/or area (a cluster) is usually the first indicator of a potential outbreak. However, the addition of a new PFGE profile to the database requires that an infected individual have symptoms severe enough to cause the patient to seek medical attention, and a physician to culture and identify the bacteria,

and to report the isolate to a PulseNet-participating laboratory for PFGE profiling. As a result of the multiplicity of steps in this process, many cases of bacteria-induced gastroenteritis go unreported.

Recent outbreaks of foodborne illnesses caused by STEC and *SE* have been correlated to a variety of food sources. Traditional sources commonly associated with *SE* and *E. coli* contamination include poultry, beef, eggs, pork, and fresh produce (Benenson and Chin, 1995; Griffin and Tauxe, 1991). Both *SE* and STEC have been associated recently with major foodborne outbreaks involving diverse foods such as cantaloupe, alfalfa sprouts, red and black pepper, turkey burgers, bologna, hazelnuts, and cheese (Maki, 2009; CDC, 2011a). Alfalfa sprouts are among the foods most commonly contaminated by *SE* and STEC; comprising the source of at least 30 separate illness outbreaks since detailed surveillance began in 1996. The ability to accurately identify and characterize the primary source of contamination in foodborne illness outbreaks allows greater insight into the natural history, epidemiology, and evolution of foodborne pathogens, all of which facilitate efforts to improve public health (Foley *et al.,* 2007).

Foodborne illness outbreak investigations

According to the CDC (2011b; 2011c), foodborne outbreak investigations consist of 7 key steps: 1) detecting a possible outbreak, 2) defining and finding cases, 3) generating hypotheses about likely sources, 4) testing the hypotheses, 5) finding the point of contamination and the source of the food, 6) controlling the outbreak, and 7) determining when the outbreak is over. None of these steps can be accomplished without robust capabilities to detect and discriminate the causative strains of pathogens. The

cornerstone of effective epidemiological surveillance and outbreak identification for foodborne pathogens lies in accurate discrimination among related isolates. A high level of pathogen strain discrimination allows better identification and monitoring of outbreaks and provides a higher probability of locating the source of an outbreak (Jones *et al.,* 2004).

Foodborne illness outbreak traceback in the U.S. is usually a collaborative effort involving the CDC, the USDA, and/or the FDA, depending on the type of food implicated. The investigation is usually initiated with detailed interviews of the infected individuals to determine candidate food sources that may have been contaminated. Once a potential source is identified, the entire pathway of the food production, transportation, and preparation is investigated for possible points of contamination, and samples from each node of this process may be assessed for the presence and type of the pathogen. If the pathogen is detected in any of these samples, the isolates are compared to the outbreak strain. Based upon the specific circumstances and progress of the investigation, the investigators may notify the general public of the outbreak and of a certain food or food group that may be contaminated (CDC, 2011b; 2011c). It is optimal if the exact source and means of contamination can be identified, but this is not always possible.

The primary means by which an outbreak can be traced back to the source is by pathogen strain discrimination, which allows epidemiologists to correlate clinical isolates with food or environmental isolates. Strain discrimination among foodborne pathogenic bacteria can be assessed by multiple phenotypic and genotypic methods, the effectiveness of which relies on three key factors—discriminatory power, typeability, and reproducibility. Discriminatory power refers to the ability of a method to separate

nonrelated strains by some measureable characteristic, typeability refers to the ability of a method to produce an interpretable result for all strains typed, and reproducibility refers to the ability of a method to produce consistent results with successive testing (Bush and Nitschko, 1999). PFGE is widely known as the bacterial subtyping "gold standard" and in most cases provides the discriminatory power, typeability, and reproducibility necessary for efficient and effective outbreak traceback (Gerner-Smidt *et al.,* 2006). However, the need for improvement exists, especially for discrimination between sporadic and outbreak cases or when an outbreak strain has a very common PFGE profile, as demonstrated by a recent SE-induced illness outbreak involving salami.

Beginning in July 2009, an increased occurrence in the number of *Salmonella* Montevideo profiles uploaded to PulseNet was identified and an outbreak investigation was initiated. Because the PFGE profile of the outbreak strain of *S.* Montevideo was the most common PFGE profile for this serovar in the database, distinguishing outbreak cases from sporadic cases was difficult. As a result, it was difficult for investigators to accurately identify the contaminated food source. Initially, sliced salami in variety packs produced by a particular manufacturer were implicated and recalled. Subsequent investigations then indicated that red and black pepper used for seasoning the salami was the real culprit and *Salmonella* isolates obtained from the peppers more closely resembled the outbreak strain than any other samples. Ultimately, another subtyping technique, single nucleotide polymorphism (SNP) analysis, was used by the FDA to confirm the source to be black pepper and crushed red pepper (Lienau *et al.,* 2011).

The salami/pepper-associated illness outbreak, causing 272 illnesses and 52 hospitalizations, was significant in that it highlighted the need for a pathogen subtyping

method with a higher level of discriminatory power than currently possible with PFGE. Since PFGE was not able to provide the level of pathogen strain discrimination necessary to trace back this outbreak, investigators were forced to resort to a more costly and time consuming method (SNP analysis) and the resulting delay may have been to blame for additional illnesses (CDC, 2012).

Another major limitation of the use of PFGE during investigations of foodborne illness outbreaks is time; two days are usually required to complete the entire process. Even if PFGE did provide the necessary discriminatory power for successful foodborne pathogen traceback, the amount of time required could mean that more illnesses may occur than if a more rapid strain discrimination method was used. The speed limitation of PFGE was emphasized by another recent foodborne illness outbreak involving STEC O104 and fenugreek sprouts.

In the summer of 2011, a severe and highly publicized outbreak of foodborne illness began in Germany and later affected individuals in other European countries and the U.S. (Scheutz *et al.,* 2011). Initially, a clear correlation with a particular food source was particularly difficult to identify (Park, 2011). The outbreak strain previously was only rarely associated with HUS and Shiga-toxin production and was found to carry genes previously associated exclusively with EAEC. As a result, this strain showed greater virulence and a higher percentage of HUS development than is typical of other STEC (Mellmann *et al.,* 2011). Although next-generation whole-genome sequencing of the outbreak strain was completed early in the outbreak investigation, this information was not sufficient to pinpoint the source of the contamination until late in the investigation due to the time required for analysis (Frank *et al.,* 2011; Park, 2011).

Overall, 4,137 illnesses were confirmed, including 830 cases of HUS and 46 deaths (Wu *et al.,* 2011; Mellmann *et al.,* 2011). In the case of emerging pathogens such as STEC O104:H4, rapid strain discrimination methods are critical. Although PFGE data greatly aided the investigation process, it is likely that had a more rapid and more discriminatory subtyping method been available at the first stages of the investigation, the outbreak source could have been more expediently identified.

Foodborne pathogens, biosecurity, and forensic microbiology

Outbreaks of foodborne illness have the potential to be naturally, unintentionally, and deliberately incited. Foodborne bacterial pathogens have been and remain attractive agents for weaponization for the purpose of bioterrorism and biocrime (Christopher *et al.,* 1997). Biosafety in Microbiological and Biomedical Laboratories (BMBL) guidelines established by the CDC list nontyphoidal *SE* and STEC as documented hazards to laboratory personnel and specify that their handling must be done in strict compliance with BSL-2 practices, containment equipment, and facilities (CDC, 2009). However, since these pathogens are not select agents and USDA Animal and Plant Health Inspection Service (APHIS) permits are generally not required for their transport, knowledgeable actors wishing to inflict harm may be able to obtain these bacteria with relative ease. Food supplies are often transported widely and are largely unguarded, allowing ease of access and difficulty in identifying the point of contamination. Additionally, outbreaks of foodborne illness often cause general fear of specific food items and panic by the population, as well as potentially severe, localized economic repercussions (CDC, 2011b).

Salmonella has already been used as a bioweapon on several occasions. The most notable case was the first incidence of known bioterrorism enacted on American soil and is one of the largest known acts of bioterrorism to date. In 1984, members of the Oregon based Rajneeshee cult, under the direction of the religious guru Bhagwan Shree Rajneesh, contaminated 10 restaurant salad bars in The Dalles, Oregon with *Salmonella* Typhimurium with the intent of influencing a local election. Local authorities reported 751 confirmed illnesses and 45 hospitalizations. At the time, this outbreak was thought to be of natural causes. Following the confessions a few individuals involved in the attack a year later, clinical isolates of the outbreak strain were found to be identical to isolates obtained from a laboratory in the commune (Török *et al.,* 1997; Carus, 2005, 1999). A robust surveillance system and better strain discrimination techniques could have allowed this outbreak to be detected more rapidly and may have allowed it to be identified as intentional much sooner.

During and after World War II, several countries, including the U.S, invested in programs aimed at the research and development of biological weapons (Christopher *et al*., 1997). One of the most extensive bioweapons R&D programs was supported by the Imperial Japanese Army and included mass production of *Salmonella enterica* (Harris, 1992). The infamous biological warfare research facility code-named Unit 731 was located in occupied China and was the center of the Japanese bioweapons development program. Extensive testing of pathogen effectiveness and dissemination techniques using animal and human subjects was carried out in this facility (Christopher *et al.,* 1997; Klietmann and Ruoff, 2001). *Salmonella* Typhi, along with many other pathogens, reportedly was used to contaminate water sources and food supplies and may have been

dispersed in aerosol form by aerial spraying and the dropping of small, specially designed bombs (Klietmann & Ruoff, 2001). Multiple outbreaks of typhoid fever, as well as several other diseases, were attributed to this program (Harris, 1992) and may have been major contributors in the deaths of up to 580,000 Chinese civilians (Barenblatt, 2004).

Determining whether an outbreak was naturally or deliberately incited is a critical first step when investigating a potential case of criminal bioweapon use. Answering this question will allow investigators to decide if a crime has been committed and is a prerequisite for attribution (Budowle *et al.,* 2005). The ability to differentiate between natural and deliberate outbreaks, however, requires the establishment of baselines of illness occurrence within a population based on thorough surveillance. The need for baselines of illness occurrence is especially important for commonly occurring outbreaks such as those caused by *Salmonella* and STEC. Ultimately, determining whether outbreaks are natural or intentional requires evidence unique to each situation, much like any other type of forensic investigation. Evidence of this nature often allows investigators to correlate separate items of evidence, all with the goal of attribution (Budowle *et al.,* 2005). In the case of foodborne outbreaks, the acquisition of unique evidence requires capabilities allowing highly sensitive pathogen detection and the highest possible levels of strain discrimination.

Microbial forensics. Budowle *et al.* (2003) define microbial forensics as a scientific discipline dedicated to analyzing evidence from a bioterrorism act, biocrime, or inadvertent microorganism/toxin release for attribution purposes. At the heart of microbial forensics is a deep and thorough understanding of the microbes used in such an

event in an attempt to identify individual characteristics that will allow attribution (Budowle *et al.,* 2005). A critical aspect of microbial forensics is the potential and probability that evidence and methods of evidence analysis will be presented in courts of law. As a result, exceptionally high levels of stringency, validation, and resolution are necessary for confident attribution (Budowle, 2004). According to the *Daubert* standard of 1993 regarding the admissibility of evidence in courts, expert testimony must be reliable, relevant, and based on scientific knowledge obtained via proper use of the scientific method. Additionally, methods used to analyze evidence must be testable, subjected to peer review, have an established error rate, have specified standards, and be generally accepted by the scientific community. Forensic methods used for the analysis of microbes implicated in biocrime and bioterrorism must meet or exceed these requirements (Budowle, 2004). The methods of critical importance to microbial forensics are high confidence level molecular typing of microbes associated with biocrime or bioterrorism events (Fletcher *et al.,* 2006). These methods allow the individual characteristic evidence necessary for successful attribution.

One very important component of agricultural biosecurity is food biosecurity. A comprehensive "Food Protection Plan" developed by the FDA mentioned that food must be considered as a potential vehicle for intentional contamination (FDA, Food Protection Plan, 2007). From the perspective of plant pathology, food biosecurity includes the study of the interactions, survival, and proliferation of human pathogens on plants and how they interact with the natural flora on and within plants. Much research is needed to elucidate these relationships in an attempt to better understand and prevent unnecessary human infection. Many outbreaks of foodborne illness are associated with the

consumption of plants or plant products (CDC, 2011). When such outbreaks occur it is essential to rapidly identify both the responsible pathogen and the implicated food as specifically as possible. Such typing must allow discrimination between infections caused by outbreak related strains and non-outbreak related strains. In the case of intentionally released pathogens, this subtyping capability must provide a confidence level that allows successful traceback and criminal attribution.

Detection of foodborne pathogens

Successful detection of foodborne pathogens in a variety of backgrounds is a critical first step in studying foodborne pathogens and is the key to prevention and identification of the health and economic burdens of foodborne illness (Velusamy *et al.,* 2009). Pathogen detection methods can be broadly separated into conventional methods and molecular-based methods. While the success and accuracy of conventional detection methods has been well documented and demonstrated, conventional detection methods primarily suffer from a lack of speed (Malorny *et al.,* 2003a). Rapid detection of foodborne pathogens is of critical importance when human illness is involved.

Conventional detection methods. Traditional or conventional detection methods used for pathogenic bacteria are dependent upon phenotypic, biochemical, or immunologically based characteristics. Following a pre-enrichment step, bacteria are usually plated onto selective media and identified by specific characteristics of colony growth and morphology. Biochemically based tests differentiate bacteria based on some biochemical trait, such as carbon utilization, and immunologically based methods use

antigen-antibody interactions for detection (Velusamy *et al.,* 2010). Although conventional methods are among the most reliable and accurate techniques for foodborne pathogen detection, a major drawback of these methods is the time necessary for positive detection. The facts that pre-enrichment is often necessary, 2-3 days are usually required for initial results, and a further 7-10 days are needed for confirmation and additional tests, are significant disadvantages in foodborne illness outbreaks where time is critical (Velusamy *et al.,* 2010). An additional drawback of conventional detection methods is the inability to detect viable but non-culturable pathogen cells.

Molecular-based detection methods. Molecular-based pathogen detection methods have the potential to overcome many of the drawbacks of conventional methods. Pathogenic *E. coli* and *SE* strains were among the first pathogens for which molecular detection methods were developed (Nataro and Kaper, 1998). The polymerase chain reaction (PCR) is the foundation of many molecular-based detection methods, fulfilling most of the requirements for widespread use and is becoming one of the primary means by which foodborne pathogens are detected in clinical and environmental samples (Malorny *et al.,* 2003a). PCR generally requires multiple cycles of a template DNA denaturing step, a primer to template DNA annealing step, and a primer elongation step. PCR allows amplification of a very specific DNA sequence to a level that can be easily visualized in agarose or polyacrylamide gels following electrophoresis. PCR-based detection assays mainly target genes that are expressed exclusively in a single genus or species. Virulence associated genes, antibiotic resistance genes, and genes acquired by

horizontal gene transfer are targeted most often. The presence or absence of an amplified DNA fragment allows detection of a specific pathogen.

Many variations of the traditional PCR method have been developed, mainly with the goal of increasing assay speed and sensitivity. Reducing the limit of detection allows pathogens to be detected at lower titers in food and clinical isolates, a feature very useful for enteric bacteria that have low infective doses (DuPont *et al.,* 1989). Multiplex PCR allows multiple primer sets to amplify fragments of multiple DNA sequences in a single assay. From the perspective of pathogen detection, multiplex PCR allows multiple pathogens to be detected simultaneously and therefore reduces the time and materials necessary for similar results using singleplex PCR procedures.

Other PCR detection assays have been developed with the intention of overcoming selective isolation and identification problems associated with the presence of normal background flora and other cellular components in isolates (Clark, 1980). These include both clinical (Stone *et al.,* 1994) and food isolates including but not limited to chicken (Soumet *et al.,* 1994), shellfish (Bej *et al.,* 1994), swine (Nucera *et al.,* 2006; Nowak *et al.,* 2007), raw oysters (Vazquez-Novelle *et al.,* 2005), alfalfa seed (Liao and Schollenberger, 2003), seeded alfalfa sprouts (Johnston *et al.,* 2005), cantaloupe (Espinoza-Medina *et al.,* 2006), tomatoes (Guo *et al.,* 2000), animal feeds (Lofstrom *et al.,* 2004), and other complex food matrices (Lampel *et al.,* 2000).

Quantitative real-time PCR (qPCR) allows visualization of the product amplification in real time as the PCR is underway. qPCR procedures also allow colorimetric or chemiluminescent molecules to be used to visualize the amplification and allows large numbers of samples to be run simultaneously while maintaining high

sensitivity and remaining relatively easy to perform (Wahlberg *et al.,* 1990). Several researchers have worked to optimize qPCR detection assays with modifications aimed at increasing sensitivity or reducing the assay run time by decreasing or eliminating enrichment steps. Examples include examining the effects of selective vs. nonselective enrichment (Eyigor *et al.,* 2002), eliminating enrichment (Wolffs *et al.,* 2006), using buoyant density gradient centrifugation to separate bacterial cells and reduce the need for enrichment (Fukushima *et al.,* 2007), overcoming background inhibition in meats (Wang *et al.,* 2004; Notzon *et al.* 2006) and milk (Van Kessel *et al*., 2003; Karns *et al.,* 2005; Nam *et al*., 2005), and using immuno-magnetic bead qPCR assays (Mercanoglu and Griffiths, 2005; Notzon *et al.,* 2006) to name a few. qPCR allows levels of sensitivity and specificity unparalleled by many other detection methods (Swaminathan and Feng, 1994).

Nested PCR assays, in which a second primer set amplifies a fragment contained within a larger amplified fragment, have been found to increase detection sensitivity (Rychlik *et al.,* 1999; Waage *et al.,* 1999). Other PCR variations that reduce the limit of detection include the use of immuno-magnetic beads (Fluit *et al.,* 1993), magnetic bead separation combined with slot blots (Li *et al.,* 2000), and adding an enrichment step prior to PCR (Myint *et al.,* 2006). Additionally, 16S ribosomal DNA (rDNA) has been used to reduce the limit of detection since rDNA is highly expressed in most cells, thus providing more template DNA per cell for amplification (Lin *et al.* 2004).

Loop-mediated isothermal amplification (LAMP), a newly developed PCR-based method that allows all PCR steps to be carried out at a single temperature and with a very short assay time, shows great promise in detection of foodborne pathogens, especially in

field applications (Hara-Kudo *et al.,* 2005; Okamura *et al.,* 2008; Chen *et al.,* 2011). In some cases, LAMP has outperformed qPCR in sensitivity (Parida *et al.,* 2005). Although not PCR-based, microarrays are being used with increasing frequency for pathogen detection and bring new levels of efficiency and reliability with high throughput (Mumford *et al.,* 2006). Additionally, there is much interest in the further development and refinement of biosensors for rapid detection of foodborne pathogens (Arora *et al.,* 2011; Velusamy *et al.,* 2010).

Small changes in PCR methods can sometimes lead to significant improvements in sensitivity or specificity. Afonina *et al.* (2007) found that by adding short, noncomplementary, AT-rich nucleotide sequences to the 5' end of PCR primers, the sensitivity of real-time PCR was increased, especially when applied to templates otherwise difficult to amplify such as viral and bisulfite-treated DNA. These short 5' additions to PCR primers have been used to add restriction sites (Espelund and Jacobsen, 1992) or universal detection sites (Li *et al.,* 2006) to PCR products in addition to reducing the number of errors when sequencing short PCR products (Binladen *et al.*, 2007).

PCR-based detection of *SE* and STEC. One of the most common and most validated, highly specific primer sets used in conventional PCR assays for detection of *SE* was developed by Rahn *et al.* (1992), amplifying a 284 bp fragment of the *invA* gene (Aabo *et al.* 1993). This primer set has been used by numerous investigators for both detection and enumeration, and is often used as a standard for comparison when new primer sets are developed (Levin, 2009; Malorny *et al.,* 2003b). Other extensively-used primer sets include those targeting the *Salmonella* origin of replication (Widjojoatmodjo

et al., 1991; Malorny *et al.,* 2003b), the *his* gene encoded within a histidine transport operon (Cohen *et al.,* 1993), the *afgA* gene encoding a component of thin aggregative fimbriae (Doran *et al.,* 1993; Craciunas *et al.,* 2010), the *fimA* gene encoding the major fimbrial unit of aggregative type 1 fimbriae (Cohen *et al.,* 1996), the *ompC* gene encoding the outer membrane protein C (Kwang *et al.,* 1996), the *iroB* gene encoding a protein involved in iron regulation (Baumler *et al.,* 1997), and the *sirA* and *hilA* genes encoding positive regulators of other *Salmonella* invasive genes (Guo *et al.,* 2000), among others.

PCR-based STEC detection is primarily focused on virulence associated genes (Nataro and Kaper, 1998). Specifically, Shiga toxin 1 (*stx1*), Shiga toxin 2 (*stx2*), and intimin (*eae*), an intestinal adherence factor, are the primary targets used for PCR detection of STEC (Nataro and Kaper, 1998). All three of these genes were recently incorporated into a single multiplex PCR assay developed by Fratamico *et al.* (2011) which was very specific and sensitive for STEC detection. Another useful multiplex PCR assay developed recently targets different parts of the O-antigen flippase genes (*wzx*) of the 6 most commonly isolated STEC serogroups in the U.S. and allows simultaneous identification of all 6 serogroups. It was found to be very specific, with clearly distinguishable agarose gel bands for each serogroup (DebRoy *et al.,* 2011).

Foodborne pathogen detection needs. An increasingly globalized food supply with the ability to transport food products around the world in a matter of hours necessitates the development of the most robust, sensitive, and rapid pathogen detection methodologies possible. Rapid and sensitive identification of the foodborne pathogens

causing human illness are important in outbreak investigations and help discriminate true positives from the inevitable false positives and reduce the number of false negatives (Gracias and McKillip, 2004).

The major limitation associated with conventional PCR detection methods relates to the time required to isolate and culture pathogens, extract DNA, perform the PCR assay, and visualize the results via agarose gel electrophoresis. This process is generally too time-consuming for widespread standardized use and the lack of the capability for automation allows inter-laboratory variation that may make the interpretation of results difficult (Malorny *et al.,* 2003a). Additionally, the process of visualization on agarose gels is not suited for high throughput detection in food quality-control laboratories (Swaminathan and Feng, 1994).

However, PCR-based methods for detecting foodborne pathogens are still among the most sensitive, rapid, and robust detection methods currently available for widespread use. The primary focus should be on improving the basic PCR procedures in innovative and creative ways that allow fewer target pathogens to be detected in less time while minimizing the effects of PCR inhibitors. Accomplishing this goal may be possible even with small PCR modifications, such as the addition of AT-rich non-complementary nucleotide sequences to the 5' end of PCR primers (Afonina *et al.,* 2007).

Discrimination of foodborne pathogens

Methods used for discrimination of foodborne pathogens can target either phenotypic or genotypic characteristics (Foley *et al.,* 2007). Phenotypic typing methods are based on patterns of gene expression that may or may not be observed in a given

strain, while genotypic typing methods are based on the DNA sequence of the bacteria and are independent of gene expression (Arbeit, 1995).

Serotyping, the traditional method of differentiating isolates of foodborne bacteria, is the basis of the *Salmonella* and *E. coli* serovar classification system established by Kauffman (1966) and White (1926). Other phenotypic methods for strain discrimination are based on antimicrobial resistance patterns, susceptibility to bacteriophages (phage typing), or charge and migration variations of key enzymes in an electric field (multi-locus enzyme electrophoresis) (Foley *et al.,* 2007). However, due to limited numbers of O and H antibodies and phages, in addition to widespread antimicrobial resistance, phenotypic methods often lack the discriminatory power necessary to differentiate closely related strains of clinically important bacteria (CDC, 2000; USDA, 2000).

Genotypic typing methods have progressed from those that require no former knowledge of the genomic DNA sequence of a particular pathogen to those that require extensive prior knowledge of it. As more and more bacterial genomes are being sequenced, a better understanding of microbial phylogenies and the genotypic differences between similar strains is likely to increase (Wu *et al.,* 2009). Current genotypic typing methods incorporate restriction enzyme digestion, PCR, and sequencing (Foley *et al.,* 2007). Those that require no prior knowledge of the genome sequence include PFGE, plasmid analysis, ribotyping, amplified fragment length polymorphism analysis (AFLP), and randomly amplified polymorphic DNA (RAPD) analysis. Methods that require prior knowledge of the genomic DNA sequence include multi-locus sequence typing (MLST),

single nucleotide polymorphism (SNP) analysis, and multiple-locus variable number tandem repeat (VNTR) analysis (MLVA).

Plasmid analysis. Plasmid analysis relies on the assessment of variations in number, size, content, and the presence of plasmids in bacterial isolates. The method requires that plasmid DNA be isolated separately from chromosomal DNA (Birnboim and Doly, 1979; Kado and Liu, 1981; Foley *et al.,* 2007). Following isolation, plasmids (whole or restricted) are electrophoretically separated. The resulting plasmid profile can be used for bacterial strain typing (Nauerby *et al.,* 2000). Plasmid profiling with *Salmonella* and STEC has limited effectiveness since some strains lack plasmids, gain and loss of plasmids can occur frequently, and low copy numbers might hinder isolation (Hoszowski and Wasyl, 2001; Liebana *et al.,* 2001; Kumao *et al.,* 2002). Thus, plasmid profiling is usually used only for *Salmonella* and STEC typing in short-term epidemiological studies (reviewed by Foley *et al.*, 2007).

Ribotyping. Ribotyping, a form of restriction fragment length polymorphism (RFLP) analysis, allows bacterial discrimination based on variations in the number and locations of gene sequences of ribosomal RNA (rRNA) digested with restriction enzymes and visualized by blotting and DNA hybridization following gel electrophoresis. As described by Bouchet *et al.* (2008), the DNA polymorphisms that allow strain discrimination are usually a result of variations on the housekeeping genes flanking chromosomal rRNA gene sequences, rather than variation in the rRNA gene itself. Although ribotyping is highly reproducible, has the potential for automation, and produces easily interpreted results, a restricted number of rRNA genes in some serotypes limit the overall potential of this method (Bouchet *et al.,* 2008). Ribotyping of *Salmonella*

serotypes might not discriminate among unrelated isolates within a serotype, a major limiting factor of this method when applied to outbreak surveillance (Kumao *et al.*, 2002). However, since this method identifies minor polymorphisms in the genes flanking the highly conserved rRNA genes, ribotyping can be of great value for elucidating the evolution of closely related strains (Bouchet *et al.,* 2008).

Randomly amplified polymorphic DNA analysis (RAPD). In RAPD analysis, short, random PCR primers that anneal to and amplify random fragments of genomic DNA are used to generate a genetic profile (Busch and Nitschko, 1999). These methods require no prior knowledge of the genome sequence, only a small amount of DNA is required, results can be obtained relatively rapidly, and more sets of primers can be used to increase discrimination. Disadvantages include poor reproducibility due to the use of nonspecific primers and possible minor variations in PCR reaction conditions and reactant concentrations (Foley *et al.,* 2007). Thus, RAPD analysis has limited value for *Salmonella* and STEC strain discrimination.

Amplified fragment length polymorphism analysis (AFLP). AFLP identifies polymorphisms within restricted and PCR-amplified DNA fragments. Following restriction digestion, DNA fragments with sequences complementary to the restriction sites are ligated to the free DNA ends. These "linker" sequences subsequently serve as targets for PCR primers, and amplified fragment profiles correspond to different strains. Since PCR primers are based on the linker sequences, no previous knowledge of the genomic sequence is necessary. Overall, AFLP is usually highly reproducible (Janssen *et al.,* 1996; Reche *et al.,* 2003), allows automation when labeled PCR primers are used (Tamada *et al.,* 2001; Deswai *et al.,* 2001), requires only a small amount of genomic

DNA due to PCR amplification (Vos *et al.,* 1995), and provides relatively thorough coverage of the entire genome (Foley *et al.,* 2007). When used to discriminate *Salmonella* and STEC strains it provides good differentiation with reasonable accuracy (Lindstedt *et al.,* 2000; Scott *et al.,* 2001; Heir *et al.,* 2000). However, potential problems with this method include the necessity to produce small enough restriction fragments for PCR amplification and electrophoretic separation and also the necessity of an automated DNA sequencer for consistent results (Tamada *et al.,* 2001). Some laboratories might not have the resources available for such equipment and the use of different sequencing platforms can hinder inter-laboratory comparison (Fry *et al.,* 2005). However, with careful standardization and validation, AFLP analysis shows promise for *Salmonella* and STEC strain discrimination.

Multiple locus sequence typing (MLST). MLST is based on polymorphisms in multiple, highly conserved genes. Genetic differences in five to ten highly conserved housekeeping genes encoding basic cellular functions provide variation for strain discrimination. Target genes are amplified by PCR and then sequenced to identify and compare nucleotide polymorphisms within the same genes in different strains. MLST can provide very high discriminatory power for many species of bacteria and is useful in outbreak epidemiology (Urwin and Maiden, 2003). In the case of *Salmonella* and STEC, MLST is attractive because it analyzes variation in the actual DNA sequence rather than the size of amplified fragments, has high inter-laboratory reproducibility, and is conducive to electronic data sharing, all of which add to the accuracy and usefulness of this method while allowing it to be unambiguous and portable (Maiden *et al.,* 1998; Cooper and Feil, 2004). However, due to small degrees of genomic variation in many

Salmonella and STEC strains, MLST is not as useful as PFGE for strain discrimination (Fakhr *et al.,* 2005). In such cases, genes that are under higher selective pressure and exhibit more genetic variation, such as virulence genes, can be used to improve discrimination (Maiden *et al.,* 1998). Online MLST databases for multiple species of bacteria, including *Salmonella* and STEC, facilitate rapid sharing of MLST data (Enright and Spratt, 1999). A thorough knowledge of the bacterial genome is required for this method, a possible limitation in emerging outbreak analyses. Identifying genes with sufficient variation, that are also flanked by stable sequences that allowing the use of a common PCR primer set for all strains is essential (Cooper and Feil, 2004). Additionally, the requirements of PCR amplification and sequencing add time as well as equipment expense and further reduce its usefulness as a rapid and widespread strain discriminatory method for foodborne outbreak related cases.

Pulsed field gel electrophoresis (PFGE). PFGE is a molecular technique used to differentiate similar strains of bacteria based on the patterns that occur in DNA fragments separated by electrophoresis on agarose gels when bacterial genomic DNA is digested with rare cutting enzymes (Smith *et al.,* 1987). PFGE was developed originally to resolve large DNA fragments that cannot be separated by standard gel electrophoresis (Schwartz and Cantor, 1984) by using an alternating voltage gradient. By periodically changing the direction of the field, DNA fragments of varying lengths react to the change at different rates. The larger fragments take longer to align to the changing fields and thus move more slowly through the gel. Over time and constantly changing fields, the bands separate more and more. Using this method, the upper resolution limit of DNA fragment size is over 10Mb, about twice as large as the *SE* and *E. coli* genomes. As a

result, entire *SE* and *E. coli* chromosomes can be analyzed. The voltage, time intervals between angle changes, and other variables are adjusted appropriate to the species under investigation (Goering, 2010; Olson, 1989). The resulting PFGE band profile can be compared with those of other strains (Tenover *et al*., 1995). With restriction enzyme optimization for specific species, this method is highly reproducible at both the intra- and inter-laboratory level. PFGE, currently the "gold standard" bacterial strain discrimination technique, is the basis of the CDC's PulseNet database (Swaminathan *et al.,* 2001). However, PFGE is time consuming and expensive, requires a high level of skill, may yield different results among operators, may not provide optimal separation in all parts of the gel, and may yield bands that contain fragments of different size (Swaminathan and Barrett, 1995). Multiple PFGE runs using different restriction enzymes may be needed to adequately type some strains (Joyner and Kincaid, 2006; Zhao *et al.,* 2006; Brown and Keys, 2006). In some cases PFGE is not a true phylogenetic measure and some strains remain untypeable by this method.

Single nucleotide polymorphism (SNP) analysis. SNP analysis is based on single nucleotide polymorphisms between strains on a genome wide basis and potentially provides the highest level of bacterial strain discrimination (Holt *et al.,* 2008). Multiple SNP locations throughout the genome are usually identified and compared between isolates in order to evaluate the genetic variation of outbreak and non-outbreak related strains. While SNP analysis is exceptionally useful for bacterial strain discrimination, it is not currently practical for epidemiological studies of foodborne illness outbreaks where time is an important factor (Ahmed *et al.,* 2008). The equipment necessary for SNP analysis is also very costly and unobtainable by many laboratories. With new

developments in technology, the ever decreasing speed and costs of DNA sequencing, and improvements in analysis software, SNP analysis may be heavily relied upon in the future for high confidence strain discrimination of outbreak related bacteria.

Multiple-locus variable-number tandem repeat (VNTR) analysis (MLVA). MLVA is a discriminatory method based on the polymorphisms in multiple VNTR locations throughout a genome. VNTRs are segments of tandemly repeated short DNA sequences ranging from just a few base pairs to several hundred. The locations and sequences of VNTRs are usually well conserved in a species but differences in the numbers of repeated units may be highly variable among closely related strains. These regions are believed to be among the most rapidly evolving segments of DNA and might be of great value in elucidating the level of evolutionary divergence between species, subspecies, and strains. MLVA has gained popularity in the forensic science community and is commonly referred to as "DNA fingerprinting" due to its highly individualistic discriminatory capability (Jeffreys *et al.,* 1985). MLVA is performed by PCR amplification of multiple VNTR loci of known repeat length and identity followed by accurate sizing of the amplified fragments in order to determine the repeat copy numbers. Multiple VNTR loci are used in order to increase the level of discrimination since variation in some loci might not be present. Therefore, a successful MLVA assay requires careful selection of VNTR loci, good PCR primer and multiplex PCR design, and an accurate method for amplicon sizing. MLVA assays are advantageous to other methods in that they are fairly rapid, allow high discrimination, and have the potential for automation when fluorescently labeled primers are used with capillary gel electrophoresis. In the case of bacteria, MLVA has been successfully used to

discriminate epidemiologically related strains of *Francisella tularensis* (Johansson *et al.,* 2001), *Bacillus anthracis* (Keim *et al.,* 2000), *Yersinia pestis* (Klevytska *et al.,* 2001), STEC O157 (Hyytia-Trees *et al.,* 2006; Lindstedt *et al.,* 2003a, 2004a; Noller *et al.,* 2003; Keys *et al.,* 2005) and *Salmonella enterica* (Lindstedt *et al.,* 2003b, 2004b; Ramisse *et al.,* 2004; Witonski *et al.,* 2006; Cho *et al.,* 2007; Boxrud *et al.,* 2007; Malorny *et al.,* 2008), among others. Recently, MLVA assays for *S.* Typhimurim, *S.* Enteritidis, and STEC O157:H7 have been added to PulseNet International and are being used in addition to PFGE when investigating and monitoring outbreaks involving these serovars. A major limitation to developing more MLVA assays for more serovars is that in many cases it appears to be necessary to select different VNTR loci for each serotype, requiring a thorough prior knowledge of the genome sequence of each serotype. With careful selection of conserved VNTR loci, it may be possible to develop a single MLVA assay that could discriminate multiple serotypes. While MLVA has great potential in *Salmonella* and STEC genotyping, the lack of more genome sequence information may be the limiting factor for increased MLVA assay development. Despite this minor drawback, MLVA is one of the most attractive methods for *Salmonella* and STEC strain discrimination.

Foodborne pathogen discrimination needs. Rapid, high-confidence strain discrimination is of utmost importance when investigating the source of foodborne outbreaks of *Salmonella* and STEC. It is likely that the best understanding of the epidemiology and characteristics of such outbreaks cannot be obtained through a single discrimination method, but a combination of several. Phenotypic and genotypic methods

should be used in tandem for the highest level of strain discrimination (Foley *et al.,* 2007). Of these, genotypic methods are generally more rapid and discriminatory and should be carried out first. Among genotypic methods, MLVA and SNP analysis show the most promise for a rapid, high-confidence epidemiological assessment of *Salmonella* outbreaks. MLVA has been used for subtyping the most common serovars of *Salmonella enterica* (Lindstedt *et al.,* 2003b, 2004b; Ramisse *et al.,* 2004; Witonski *et al.,* 2006; Cho *et al.,* 2007; Boxrud *et al.,* 2007; Malorny *et al.,* 2008; Beranek *et al.,* 2009) and STEC O157 (Hyytia-Trees, *et al.,* 2006; Lindstedt *et al.,* 2003a, 2004a; Noller *et al.,* 2003; Keys *et al.,* 2005). Recently, a standardized approach to the use of MLVA assays has been established (Hopkins *et al.,* 2011). The effective use of MLVA, however, may be limited to a single serovar at a time and also requires a very thorough knowledge of the genome of the infecting serovar, usually requiring very expensive, specialized equipment. The development of new MLVA assays that could allow the discrimination of two or more serovars in a single assay, while still maintaining high levels of standardization and inter-laboratory reproducibility, would be of great value.

LITERATURE CITED

Aabo, S., O.F. Rasmussen, L. Roseen, P.D. Sørensen, and J.E. Olsen. (1993). *Salmonella* identification by the polymerase chain reaction. *Mol. Cell. Probes.* 7:171-178.

Abdulmawjood, A., S. Roth, and M BuÈlte. (2001). Evaluation of a PCR system for the detection of *E. coli* O157 based on *rfbE* O157 gene. The 42nd Congress on Food Hygiene; German Veterinary Society (DVG), Sept.25 - 28 in Garmisch-Partenkirchen, Germany. pp. 368-373.

Abdulmawjood, A., M. Bülte, N. Cook, S. Roth, H. Schönenbrücher, and J. Hoorfar. (2003). Toward an international standard for PCR-based detection of *Escherichia coli* O157: Part 1. Assay development and multi-center validation. *J. Microbiol. Methods.* 55: 775-786.

Afonina, I., I. Ankoudinova, A. Mills, S. Lokhov, P. Huynh, and W. Mahoney. (2007). Primers with 5′ flaps improve real-time PCR. *BioTechniques.* 43:770-774.
Ahmed, N., U. Dobrindt, J. Hacker, and S.E. Hasnain. (2008). Genomic fluidity and pathogenic bacteria: applications in diagnostics, epidemiology and intervention. *Nat. Rev. Micro*. 6:387-394.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. (1990). Basic local alignment search tool. *J. Mol. Bio.* 215:403–410.

Angulo, F.J., A.C. Voetsch, D. Vugia D, J.L. Jadler, M. Farley, C. Hedberg, P. Cieslak, D. Morse, D. Dwyer, and D.L. Swerdlow. (1998). Determining the burden of human illness from foodborne disease: CDC's Emerging Infectious Disease Program Foodborne Diseases Active Surveillance Network (FoodNet). *Vet. Clin. North Am. Food Anim. Pract*. 14:165-72.

Arora, P., A. Sindhu, N. Dilbaghi, and A. Chaudhury. (2011). Biosensors as innovative tools for the detection of food borne pathogens. *Biosen. Bioelectron.* 28:1-12.

Arbeit, R.D. (1995). Laboratory procedures for the epidemiologic analysis of microorganisms, pp 190-208. *In* P.R. Murray, E.J. Baron, M.A. Pfaller, F.C. Tenover, and R.H. Yolken (eds.), Manual of Clinical Microbiology. ASM Press, Washington, D.C.

Barenblatt, D. (2004). A plague upon humanity. Harper, New York, NY.

Baumler, A.J., F. Heffron, and R. Reissbrodt. (1997). Rapid detection of *Salmonella enterica* with primers specific for *iroB*. *J. Clin. Microbiol.* 35:1224–1230.

Bej, A.K., M.H. Mahbubani, M.J. Boyce and R.M. Atlas (1994). Detection of *Salmonella* spp. in oysters by PCR. *Appl. Environ. Microbiol.* 60:368–373.

Benenson, A.S. and J. Chin. (1995). Control of Communicable Diseases Manual. American Public Health Association. Washington, D.C.

Beranek A, C. Mikula, P. Rabold, D. Arnhold, C. Berghold, I. Lederer, F. Allerberger, and C. Kornschober. (2009). Multiple-locus variable-number tandem repeat analysis for subtyping of *Salmonella enterica* subsp. *enterica* serovar Enteritidis. *Int. J. Med. Microbiol*. 299:43-51.

Bezuidt, O., R. Pierneef, K. Mncube, G. Lima-Mendez, and O.N. Reva. (2011). Mainstreams of horizontal gene exchange in Enterobacteria: Consideration of the outbreak of enterohemorrhagic *E. coli* O104:H4 in Germany in 2011. *PLoS ONE.* 6(10):e25702.

Birnboim, H.C. and J. Doly. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7:1513–1523.

Binladen, J., M.T. Gilbert, P.F. Campos, and E. Willerslev. (2007). 5′-tailed sequencing primers improve sequencing quality of PCR products. *BioTechniques.* 42:174-176.

Bouchet, V., H. Huot, and R. Goldstein. (2008). Molecular genetic basis of ribotyping. *Clin. Microbiol. Rev.* 21:262-273.

Boxrud, D., K. Pederson-Gulrud, J. Wotton, C. Medus, E. Lyszkowicz, J. Besser, and J.M. Bartkus. (2007). Comparison of multiple-locus variable-number tandem repeat analysis, pulsed-field gel electrophoresis, and phage typing for subtype analysis of *Salmonella enterica* serotype Enteritidis. *J. Clin. Microbiol.* 45:536-543.

Breeze, R.G. (2005). Infectious diseases: Not just a health matter anymore, pp. 27-39. *In* R.G. Breeze, B. Budowle, and S.E. Schutzer (eds.), Microbial forensics. Elsevier Academic Press, San Diego, CA.

Brenner, F.W., and A.C. McWhorter-Murlin. (1998). Identification and serotyping of *Salmonella*. Centers for Disease Control and Prevention, Atlanta, GA.

Brown, E.W. and C. Keys. (2006). Pulsed-field gel electrophoresis is a poor evolutionary indicator in *Salmonella*. Abstract from the 2006 ASM General Meeting, Orlando, FL.

Budowle, B. (2004). Genetics and attribution issues that confront the microbial forensics field. *Forensic Sci. Int*. 146(Suppl.):S185–S188.

Budowle, B., J. Burans, R.G. Breeze, M.R. Wilson, and R. Chakraborty. (2005). Microbial forensics, pp. 1–26. *In* R. G. Breeze, B. Budowle, and S.E. Schutzer (eds.), Microbial forensics. Elsevier Academic Press, San Diego, CA.

Budowle, B., S.E. Schutzer, A. Einseln, L.C. Kelley, A.C. Walsh, J.A. Smith, B.L. Marrone, J. Robertson, and J. Campos. (2003). Public health. Building microbial forensics as a response to bioterrorism. *Science.* 301:1852–1853.

Busch, U. and H. Nitschko. (1999). Methods for the differentiation of microorganisms. *J. Chromatogr. B. Biomed. Sci. Appl.* 722:263–278.

Buzby, J.C., and T. Roberts. (2009). The economics of enteric infections: Human foodborne disease costs. *Gastroenterology.* 136:1851-1862.

Carus, W.S. (1999). Bioterrorism and biocrimes. The illicit use of biological agents in the 20th century. Center for Counterproliferation Research, National Defense University, Washington, DC.

Carus, W.S. (2005). Rajneeshees. Encyclopedia of Bioterrorism Defense. John Wiley & Sons, Inc.

Celum, C.L., R.E. Chaisson, G.W. Rutherford, J.L. Barnhart, and D.F. Echenberg. (1987). Incidence of salmonellosis in patients with AIDS. *J. Infect. Dis.* 156:998-1002

Centers for Disease Control and Prevention. (1994). Addressing emerging infectious disease threats: A prevention strategy for the United States. U.S. Department of Health and Human Services, Atlanta, GA.

Centers for Disease Control and Prevention. (1997). Foodborne Diseases Active Surveillance Network (FoodNet). *Emerg. Infect. Dis.* 3:581–583.

Centers for Disease Control and Prevention. (2000). National Antimicrobial Resistance Monitoring System for Enteric Bacteria 1999 Final Report. Centers for Disease Control and Prevention, Atlanta, GA.

Centers for Disease Control and Prevention. (2008). Salmonella Surveillance: Annual Summary, 2006. US Department of Health and Human Services, Atlanta, GA.

Centers for Disease Control and Prevention. (2009). Biosafety in Microbiological and Biomedical Laboratories (BMBL) 5th Edition. pp. 151-155.

Centers for Disease Control and Prevention. (2011a). Multistate foodborne outbreaks. Available at: http://www.cdc.gov/outbreaknet/outbreaks.html. Accessed 24 August 2011.

Centers for Disease Control and Prevention. (2011b). Investigating foodborne outbreaks. Available at: http://www.cdc.gov/outbreaknet/investigations/. Accessed 24 August 2011.

Centers for Disease Control and Prevention. (2011c). Step-by-step guide to investigating foodborne outbreaks. Available at: http://www.cdc.gov/outbreaknet/investigations/investigating.html. Accessed 24 August 2011.

Centers for Disease Control and Prevention. (2012). Investigation Update: Multistate Outbreak of Human *Salmonella* Montevideo Infections. Available at: http://www.cdc.gov/salmonella/montevideo/. Accessed 27 January 2012.

Centers for Disease Control and Prevention (2012b). National Shiga toxin-producing *Escherichia coli* (STEC) Surveillance Overview. Atlanta, Georgia: US Department of Health and Human Services.

Centers for Disease Control and Prevention (2012c). Foodborne Diseases Active Surveillance Network (FoodNet). Available at: http://www.cdc.gov/foodnet/about.html. Accessed 15 October, 2012.

Chaudhury, A., G. Nath, A. Tikoo, and S.C. Sanyal. (1999). Enteropathogenicity and antimicrobial susceptibility of new *Escherichia* spp. *J. Diarrhoeal. Dis. Res.* 17:85–87.

Chen, S., F. Wang, J.C. Beaulieu, R.E. Stein, and B. Ge. (2011). Rapid detection of viable *Salmonella* in produce by coupling propidium monoazide with loop-mediated isothermal amplification (PMA-LAMP). *Appl. Environ. Microbiol.*

Christopher, L.G.W., L.T.J. Cieslak, J.A. Pavlin, and E.M. Eitzen. (1997). Biological Warfare. *JAMA—J. Am. Med. Assoc.* 278:412-417.

Clark, J.A. (1980). The influence of increasing numbers of non-indicator organisms upon the detection of indicator organisms by the membrane filter and presence-absence tests. *Can. J. Microbiol.* 26:827-32

Coburn, B., G.A. Grassl, and B.B. Finlay. (2007). *Salmonella*, the host and disease: A brief review. *Immunol. Cell. Biol*. 85:112-8.

Cohen, N.D., H.L. Neibergs, E.D. McGruder, H.W. Whitford, R.W. Behle, P.M. Ray, and B.M. Hargis. (1993). Genus-specific detection of salmonellae using the polymerase chain reaction (PCR). *J. Vet. Diagn. Invest.* 5:368–371.

Cohen, H., S. Mechanda, and W. Lin. (1996). PCR amplification of the *fimA* gene sequence of *Salmonella typhimurium,* a specific method for detection of *Salmonella* spp. *Appl. Environ. Microbiol.* 62: 4303-4308.

Cookson, B.T. and M.A. Brennan. (2001). Pro-inflammatory programmed cell death. *Trends Microbiol.* 9:113-114.

Coombes, B.K., M.E. Wickham, M. Mascarenhas, S. Gruenheid, B.B. Finlay, and M.A. Karmali. (2008). Molecular analysis as an aid to assess the public health risk of non-O157 Shiga toxin-producing *Escherichia coli* strains. *Appl. Environ. Microbiol.* 74:2153–2160.

Cooper, J.E. and E.J. Feil. (2004). Multilocus sequence typing - what is resolved? *Trends Microbiol.* 12:373-377.

Crăciunaş, C., A.L. Keul, M. Flonta, and M. Cristea, M. (2010). DNA-based diagnostic tests for *Salmonella* strains targeting *hilA, agfA, spvC* and *sef* genes. *J. Environ. Manag.* 95:S15-S18.

Desai, M., E.J. Threlfall, and J. Stanley. (2001). Fluorescent amplified-fragment length polymorphism subtyping of the *Salmonella* enterica serovar Enteritidis phage type 4 clone complex. *J. Clin. Microbiol*. 39:201–206.

DebRoy, C., E. Roberts, A.M. Valadez, E.G. Dudley, and C.N. Cutter. (2011). Detection of Shiga toxin–producing *Escherichia coli* O26, O45, O103, O111, O113, O121, O145, and O157 Serogroups by multiplex polymerase chain reaction of the *wzx* gene of the O-antigen gene cluster. *Foodborne Pathog. Dis*. 8: 651-652.

Doran, J.L., S.K. Collinson, J. Burian, G. Sarlós, E.C.D. Todd, C.K. Munro, C.M. Kay, P.A. Banser, P.I. Peterkin, and W.W. Kay. (1993). DNA-based diagnostic tests for *Salmonella* species targeting *agfA,* the structural gene for thin, aggregative fimbriae. *J. Clin. Microbio*l. 31:2263-2273

DuPont, H.L., M.M. Levine, R.B. Hornick, and S.B. Formal. (1989). Inoculum size in shigellosis and implication for expected mode of transmission. *J. Infect. Dis.* 159:26-28

Edwards, P.R., and W.H. Ewing. 1972. Identification of Enterobacteriaceae. Burgess Publishing Co, Minneapolis, Minn.

Enright, M. C., and B.G. Spratt. (1999). Multilocus sequence typing. *Trends Microbiol.* 7: 482-487.

Espelund, M. and K.S. Jacobsen. (1992). Cloning and direct sequencing of plant promoters using primer-adapter mediated PCR on DNA coupled to a magnetic solid phase. *BioTechniques.* 13:74-81.

Espinoza-Medina, I. E., F.J. Rodrguez-Leyva, I. Vargas-Arispuro, M.A. Islas-Osuna, E. Acedo-Felix, and M.A. Martinez-Tellez. (2006). PCR identification of *Salmonella*:

Potential contamination sources from production and postharvest handling of cantaloupes. *J. Food Prot.* 69:1422-1425.

Eyigor, A., K.T. Carli, and C.B. Unal. (2002). Implementation of real-time PCR to tetrathionate broth enrichment step of *Salmonella* detection in poultry. *Lett. Appl. Microbiol.* 34:37-41.

Fakhr, M. K., L.K. Nolan, and C.M. Logue. (2005). Multilocus sequence typing lacks the discriminatory ability of pulsed-field gel electrophoresis for typing *Salmonella enterica* serovar Typhimurium. *J. Clin. Microbiol.* 43: 2215-2219.

Fletcher, J., C. Bender, B. Budowle, W.T. Cobb, S.E. Gold, C.A. Ishimaru, D. Luster, U. Melcher, R. Murch, H, Scherm, R.C. Seem, J.L. Sherwood, B.W. Sobral and S.A. Tolin. (2006). Plant pathogen forensics: capabilities, needs, and recommendations. *Microbiol. Molec. Biol. Rev.* 70:450-471.

Fluit, A.C., R. Torensma, M.J. Visser, C.J. Aarsman, M.J. Poppelier, B.H. Keller, P. Klapwijk, and J. Verhoef. (1993). Detection of *Listeria monocytogenes* in cheese with the magnetic immuno-polymerase chain reaction assay. *Appl. Environ. Microbiol.* 59:1289-1293.

Foley, S., S. Zhao, and R. Walker. (2007). Comparison of molecular typing methods for the differentiation of *Salmonella* foodborne pathogens. *Foodborne Pathog. Dis*. 4:253–276.

Food and Drug Administration. (2007). Food Protection Plan. Department of Health and Human Services. Maryland: U.S. Food and Drug Administration

Francis, C.L., M.N. Starnbach, and S. Falkow. (1992). Morphological and cytoskeletal changes in epithelial cells occur immediately upon interaction with *Salmonella typhimurium* grown under low-oxygen conditions. *Mol. Microbiol.* 6:3077–87

Francis, C.L., T.A. Ryan, B.D. Jones, S.L. Smith, and S. Falkow. (1993). Ruffles induced by *Salmonella* and other stimuli direct macropinocytosis of bacteria. *Nature*. 364:639-642.

Frank, C., D. Werber, J.P. Cramer, M. Askar, M. Faber, M. Heiden, and G. Krause. (2011). Epidemic profile of Shiga toxin–producing *Escherichia coli* O104:H4 outbreak in Germany — Preliminary Report. *New Engl. J. Med.* 365:1771-1780

Fratamico, P.M., L.K. Bagi, W.C. Cray, N. Narang, X. Yan, M. Medina, and Y. Liu. (2011). Detection by multiplex real-time polymerase chain reaction assays and isolation of Shiga toxin–producing *Escherichia coli* serogroups O26, O45, O103, O111, O121, and O145 in Ground Beef. *Foodborne Pathog. Dis.* 8:601-607.

Frenzen, P.D., A. Drake, F.J. Angulo, and The Emerging Infections Program FoodNet Working Group. (2005). Economic cost of illness due to *Escherichia coli* O157 infections in the United States. *J. Food Prot.* 68:2623-2630.

Fry, N.K., B. Afshar, P. Visca, D. Jonas, J. Duncan, E. Nebuloso, A. Underwood, and T.G. Harrison. (2005). Assessment of fluorescent amplified fragment length polymorphism analysis for epidemiological genotyping of Legionella pneumophila serogroup 1. *Clin. Microbiol. Infect.* 11:704–712.

Fukushima, H., K. Katsube, Y. Hata, R. Kishi, and S. Fujiwara. (2007). Rapid separation and concentration of food-borne pathogens in food samples prior to quantification by viable-cell counting and real-time PCR. *Appl. Environ. Microbiol.* 73:92-100.

Gerner-Smidt, P., K. Hise, J. Kincaid, S. Hunter, S. Rolando, E. Hyytiä-Trees, E.M. Ribot, and B. Swaminathan. (2006). PulseNet USA: a five-year update. *Foodborne Pathog. Dis.* 3:9-19.

Gilliss, D., A. Cronquist, M. Cartter, M. Tobin-D'Angelo, D. Blythe, K. Smith, S. Lathrop, G. Birkhead, P. Cieslak, J. Dunn, K.G. Holt, J.J. Guzewich, O.L. Henao, B. Mahon, P.

Griffin, R. Tauxe, and S.M. Crim. (2011). Vital signs: Incidence and trends of infection with pathogens transmitted commonly through food - Foodborne Diseases Active Surveillance Network, 10 U.S. Sites, 1996-2010. *Morbidity and Mortality Weekly Report*. 60:10

Goering, R.V. (2010). Pulsed field gel electrophoresis: A review of application and interpretation in the molecular epidemiology of infectious disease. *Infect. Genet. Evol.* 10:866-875.

Gould, L.H., L. Demma, T.F. Jones, S. Hurd, D.J Vugia, K. Smith, B. Shiferaw, S. Segler, A. Palmer, S. Zansky, and P.M Griffin. (2009). Hemolytic uremic syndrome and death in persons with *Escherichia coli* O157:H7 infection, Foodborne Diseases Active Surveillance Network Sites, 2000–2006. *Clin. Infect. Dis.* 49:1480-1485.

Gracias, K.S. and J.L. McKillip. (2004). A review of conventional detection and enumeration methods for pathogenic bacteria in food. *Can. J. Microbiol.* 50: 883-890.

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

Griffin, P.M., S.M. Ostroff, R.V. Tauxe, K.D. Greene, J.G. Wells, J.H. Lewis, and P.A. Blake. (1988). Illnesses associated with *Escherichia coli* 0157:H7 infections. *Ann. Intern. Med.* 109:705-712.

Guo, X., J. Chen, L.R. Beuchat, and R.E. Brackett. (2000). PCR detection of *Salmonella enterica* serotype Montevideo in and on raw tomatoes using primers derived from *hilA*. *Appl. Environ. Microbiol.* 66:5248-5252.

Hadler, J.L., P. Clogher, S. Hurd, Q. Phan, M. Mandour, K. Bemis, and R. Marcus. (2011). Ten-year trends and risk factors for non-O157 Shiga toxin–producing *Escherichia coli* found through Shiga toxin testing, Connecticut, 2000–2009. *Clin. Infect. Dis.* 53: 269-276.

Hara-Kudo, Y., M. Yoshino, T. Kojima, and M. Ikedo. (2005). Loopmediated isothermal amplification for the rapid detection of *Salmonella. FEMS Microbiol. Lett.* 253:155–161.

Harris, S. (1992). Japanese biological warfare research on humans: A case study on microbiology and ethics. *Ann. N.Y. Acad. Sci*. 666:21–52.

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

Heir, E., B.A. Lindstedt, T. Vardund, Y. Wasteson, and G. Kapperud. (2000). Genomic fingerprinting of Shiga toxin-producing *Escherichia coli* (STEC) strains: comparison of pulsed-field gel electrophoresis (PFGE) and fluorescent amplified-fragment-length polymorphism (FAFLP). *Epidemiol. Infect.* 125:537–548.

Hersh, D., D.M. Monack, M.R. Smith, N. Ghori, S. Falkow, and A. Zychlinsky A. (1999). The *Salmonella* invasin SipB induces macrophage apoptosis by binding to caspase-1. *Proc. Natl. Acad. Sci.* 96:2396–2401.

Hird, S., C. Stein, T. Kuchenmüller, and R. Green. (2009). Meeting report: Second annual meeting of the World Health Organization initiative to estimate the global burden of foodborne diseases. *Int. J. Food Microbiol.* 133: 210-212.

Hoffmann, S., M.B. Batz, M. and J.G. Morris Jr. (2012). Annual Cost of Illness and Quality-Adjusted Life Year Losses in the United States Due to 14 Foodborne Pathogens. *J Food Prot.* 75(7):1292-1302.

Hogue, A., P. White, J. Guard-Petter, W. Schlosser, R. Gast, E. Ebel, J. Farrar, T. Gomez, J. Madden, M. Madison, A.M. McNamara, R. Morales, D. Parham, P. Sparling, W. Sutherlin, and D. Swerdlow. (1997). Epidemiology and control of egg associated *Salmonella* Enteritidis in the United States of America. *Rev. Sci. Tech.* 16:542–53.

Holt, K.E., J. Parkhill, C.J. Mazzoni, P. Roumagnac, F.X. Weill, I. Goodhead, R. Rance, S. Baker, D.J. Maskell, J. Wain, C. Dolecek, M. Achtman, and G. Dougan. (2008). High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. *Nat. Genet*. 40:987-993.

Hopkins, K.L., T.M. Peters, E. de Pinna, J. Wain. (2011). Standardisation of multilocus variable-number tandem-repeat analysis (MLVA) for subtyping of *Salmonella enterica* serovar Enteritidis. *Euro. Surveill.* 16(32):1-11.

Hoszowski, A. and D. Wasyl. (2001). Typing of *Salmonella enterica* subsp. *enterica* serovar Mbandaka isolates. *Vet. Microbiol*. 80:139–148.

Hyytiä-Trees, E., S.C. Smole, P.A. Fields, B. Swaminathan, and E.M. Ribot. (2006). Second generation subtyping: A proposed PulseNet protocol for multiple-locus variable-number tandem repeat analysis of Shiga toxin–producing *Escherichia coli* O157 (STEC O157). *Foodborne Pathog. Dis.* 3:118-131

Janssen, P., R. Coopman, G. Huys, J. Swings, M. Bleeker, P. Vos, M. Zabeau, and K. Kersters. (1996). Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiol.* 142:1881–1893.

Jeffreys, A.J., J.F. Brookfield, and R. Semeonoff. (1985). Positive identification of an immigration test case using human DNA fingerprints. *Nature.* 317:818–819.

Johansson, A., I. Goransson, P. Larsson, and A. Sjostedt. (2001). Extensive allelic variation among *Francisella tularensis* strains in a short-sequence tandem repeat region. *J. Clin. Microbiol.* 39:3140-3146.

Johnson, K.E., C.M. Thorpe, and C.L. Sears. (2006). The emerging clinical importance of non-O157 Shiga toxin—producing *Escherichia coli. Clin. Infect. Dis.* 43: 1587-1595.

Johnston, L.M., D. Elhanafi, M. Drake, and L.A. Jaykus. (2005). A simple method for the direct detection of *Salmonella* and *Escherichia coli* O157:H7 from raw alfalfa sprouts and spent irrigation water using PCR. *J. Food Prot.* 68:2256-2263.

Jones, T.F., B. Imhoff, M. Samuel, P. Mshar, K.G. McCombs, and M. Hawkins. (2004). Limitations to successful investigation and reporting of foodborne outbreaks: An analysis of foodborne disease outbreaks in FoodNet catchment areas, 1998–1999. *Clin. Infect. Dis.* 38:S297-S302.

Joyner, M.M. and J.A. Kincaid. (2006). Significance of second enzyme data, PulseNet National *E. coli* O157:H7 database. Proceedings from the 10th Annual PulseNet Update Meeting, Miami, FL.

Kado, C.I. and S.T. Liu. (1981). Rapid procedure for detection and isolation of large and small plasmids. *J. Bacteriol.* 145:1365–1373.

Karns, J.S., J.S. Van Kessel, B.J. McCluskey, and M.L. Perdue. (2005). Prevalence of *Salmonella enterica* in bulk tank milk from US dairies as determined by polymerase chain reaction. *J. Dairy Sci*. 88: 3475-3479.

Karmali, M., M. Petric, B. Steele, and C. Lim. (1983). Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. *The Lancet.* 321: 619-620.

Karmali, M.A., M. Petric, C. Lim, P.C. Fleming, G.S. Arbus, and H. Lior. (1985). The association between hemolytic uremic syndrome and infection by Verotoxin-producing *Escherichia coli*. *J. Infect. Dis.* 151:775-782.

Karmali, M.A. (1989). Infection by Verocytotoxin-producing *Escherichia coli*. *Clin. Microbiol. Rev.* 2:15-38.

Karmali, M.A., V. Gannon, and J.M. Sargeant. (2010). Verocytotoxin-producing *Escherichia coli* (VTEC). *Vet. Microbiol.* 140:360-370.

Kauffmann, F. (1966). The bacteriology of Enterobacteriaceae. Munksgaard, Copenhagen, Denmark.

Keim, P., L.B. Price, A.M. Klevytska, K.L. Smith, J.M. Schupp, R. Okinaka, P.J. Jackson, and M.E. Hugh-Jones. (2000). Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *J. Bacteriol.,* 182: 2928-2936.

Keys, C., S. Kemper, and P. Keim. (2005). Highly diverse variable number tandem repeat loci in the *E. coli* O157:H7 and O55:H7 genomes for high-resolution molecular typing. *J. Appl. Microbiol.* 98: 928-940.

Klevytska, A.M., L.B. Price, J.M. Schupp, P.L. Worsham, J. Wong, and P. Keim. (2001). Identification and characterization of variable-number tandem repeats in the *Yersinia pestis* genome. *J. Clin. Microbiol.* 39:3179-3185.

Klietmann, W.F. and K.L. Ruoff. (2001). Bioterrorism: Implications for the clinical microbiologist. *Clin. Microbiol. Rev.* 14: 364-381.

Knutsson, R., Y. Blixt, H. Grage, E. Borch, and P. Radstrom. (2001). Evaluation of selective enrichment PCR procedures for *Yersinia enterocolitica*. *Int. J. Food Microbiol.* 73:35–46.

Kumao, T., W. Ba-Thein, and H. Hayashi. (2002). Molecular subtyping methods for detection of *Salmonella* enterica serovar Oranienburg outbreaks. *J. Clin. Microbiol.* 40:2057–2061.

Kwang, J., E.T. Littledike, and J.E. Keen. (1996). Use of the polymerase chain reaction for *Salmonella* detection. *Lett. Appl. Microbiol*. 22:46–51.

Lampel, K.A., P.A. Orlandi, and L. Kornegay. (2000). Improved template preparation for PCR-based assays for detection of food-borne bacterial pathogens. *Appl. Environ. Microbiol.* 66:4539-4542.

Levin, R. (2009). The use of molecular methods for detecting and discriminating *Salmonella* associated with foods. A review. *Food Biotech.* 23:313-367.

Li, X., N. Boudjellab, and X. Zhao. (2000). Combined PCR and slot blot assay for detection of *Salmonella* and *Listeria monocytogenes*. *Int. J. Food Microbiol.* 56: 167-177.

Li, X., Y. Huang, Y. Guan, M. Zhao, and Y. Li. (2006). Universal molecular beacon-based tracer system for real-time polymerase chain reaction. *Anal. Chem*. 78:7886-7890.

Liao, C.H. and L.M. Shollenberger. (2003). Detection of *Salmonella* by indicator agar media and PCR as affected by alfalfa seed homogenates and native bacteria. *Lett. Appl. Microbiol.* 36:152-156.

Lienau, E.K., E. Strain, C. Wang, J. Zheng, A.R. Ottesen, C.E. Keys, R. Stones. (2011). Identification of a salmonellosis outbreak by means of molecular sequencing. *New Engl. J. Med.* 364: 981-982.

Lin, C.-K., C.-L. Hung, S.-C. Hsu, C.-C. Tsai, and H.-Y. Tsen. (2004). An improved PCR primer pair based on 16S rDNA for the specific detection of *Salmonella* serovars in food samples. *J. Food Prot. 67*:1335-1343.

Lindstedt, B.A., E. Heir, T. Vardund, and G. Kapperud. (2000). Fluorescent amplified-fragment length polymorphism genotyping of *Salmonella enterica* subsp. *enterica* serovars and comparison with pulsed-field gel electrophoresis typing. *J. Clin. Microbiol*. 38:1623–1627.

Lindstedt, B.-A., E. Heir, E. Gjernes, T. Vardund, and G. Kapperud. (2003a). DNA fingerprinting of Shiga-toxin producing *Escherichia coli* O157 based on multiple-locus variable-number tandem-repeats analysis (MLVA). *Ann. Clin. Microbiol. Antimicrob. 2*:12.

Lindstedt, B., E. Heir, E. Gjernes, and G. Kapperud. (2003b). DNA fingerprinting of *Salmonella enterica* subsp. *enterica* serovar Typhimurium with emphasis on phage

type DT104 based on variable number of tandem repeat loci. *J. Clin. Microbiol.* 41:1469 - 1479.

Lindstedt, B.-A., T. Vardund, and G. Kapperud. (2004a). Multiple-locus variable-number tandem-repeats analysis of *Escherichia coli* O157 using PCR multiplexing and multi-colored capillary electrophoresis. *J. Microbiol. Meth.* 58:213-222.

Lindstedt, B., T. Vardund, L. Aas, and G. Kapperud, G. (2004b). Multiple-locus variable-number tandem-repeats analysis of *Salmonella enterica* subsp. *enterica* serovar Typhimurium using PCR multiplexing and multicolor capillary electrophoresis. *J. Microbiol. Meth*. 59:163-172.

Lindstedt, B.-A., L.T. Brandal, L. Aas, T. Vardund, and G. Kapperud. (2007). Study of polymorphic variable-number of tandem repeats loci in the ECOR collection and in a set of pathogenic *Escherichia coli* and *Shigella* isolates for use in a genotyping assay. *J. Microbiol. Meth.* 69:197-205.

Lofstrom, C., R. Knutsson, C.E. Axelsson, and P. Radstrom. (2004). Rapid and specific detection of *Salmonella* spp. in animal feed samples by PCR after culture enrichment. *Appl. Environ. Microbiol.* 70: 69-75.

Lynch, M., J. Painter, R. Woodruff, and C. Braden. (2006). Surveillance for foodborne-disease outbreaks–United States, 1998–2002. *MMWR Surveill. Summ*. 55:1–42

Lynch, M.F., E.M. Blanton, S. Bulens, C. Polyak, J. Vojdani, J. Stevenson, and E.D. Mintz. (2009). Typhoid fever in the United States, 1999-2006. *JAMA: J. Am. Med. Assoc.* 302:859-865.

Maiden, M.C.J., J.A. Bygraves, E. Feil, G. Morelli, J.E. Russell, R. Urwin, and B.G. Spratt. (1998). Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceed. Natl. Acad. Sci.* 95:3140-3145.

Maki, D.G. (2009). Coming to grips with foodborne infection -- peanut butter, peppers, and nationwide salmonella outbreaks. *New Engl. J Med*. 360:949-953
Malorny, B., J. Hoorfar, C. Bunge and R. Helmuth. (2003b). Multicenter validation of the analytical accuracy of Salmonella PCR: towards an international standard. *Appl. Environ. Microbiol.* 69: 290–296.

Malorny, B., P.T. Tassios, P. Rådström, N. Cook, M. Wagner, and J. Hoorfar. (2003a). Standardization of diagnostic PCR for the detection of foodborne pathogens. *Int. J. Food Microbiol.* 83: 39-48.

Malorny, B., E. Junker, and R. Helmuth. (2008). Multi-locus variable-number tandem repeat analysis for outbreak studies of *Salmonella enterica* serotype Enteritidis. *BMC Microbiol.* 8:84.

Mason, J. (1994). *Salmonella* Enteritidis control programs in the United States. *Int. J. Food Microbiol.* 21:155–69.

Mathusa, E.C., Y. Chen, E. Enache, and L. Hontz. (2010). Non-O157 Shiga toxin-producing *Escherichia coli* in foods. *J. Food Protect.* 73: 1721-1736.

McClelland, M., K.E. Sanderson, J. Spieth, S.W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston, and R.K. Wilson. 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature.* 413:852–856.

McCormick, B.A., S.I. Miller, D. Carnes, and J.L. Madara. (1995). Transepithelial signaling to neutrophils by Salmonellae: A novel virulence mechanism for gastroenteritis. *Infect. Immun.* 63:2302-2309

Mellmann, A., D. Harmsen, C.A. Cummings, E.B. Zentz, S.R. Leopold, et al. (2011) Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. *PLoS ONE*. 6:e22751.

Mercanolu, B. and W. Griffiths-Mansel. (2005). Combination of immunomagnetic separation with real-time PCR for rapid detection of *Salmonella* in milk, ground beef, and alfalfa sprouts. *J. Food Prot.* 68: 557-561.

Mumford, R., N. Boonham, J. Tolinson, and I. Barker. (2006). Advances in molecular phytodiagnostics -new solutions for old problems. *Eur. J. Plant Pathol.* 116: 1–19.

Myint, M.S., Y.J. Johnson, N.L. Tablante, and R.A. Heckert. (2006). The effect of pre-enrichment protocol on the sensitivity and specificity of PCR for detection of naturally contaminated *Salmonella* in raw poultry compared to conventional culture. *Food Microbiol.* 23:599-604.

Nam, H.-M., V. Srinivasan, B.E. Gillespie, S.E. Murinda, and S.P. Oliver. (2005). Application of SYBR green real-time PCR assay for specific detection of *Salmonella* spp. in dairy farm environmental samples. *Int. J. Food Microbiol. 102*: 161-171.

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

Nauerby, B., K. Pedersen, H.H. Dietz, and M. Madsen. (2000). Comparison of Danish isolates of *Salmonella enterica* serovar Enteritidis PT9a and PT11 from

hedgehogs (*Erinaceus europaeus*) and humans by plasmid profiling and pulsed-field gel electrophoresis. *J. Clin. Microbiol*. 38:3631–3635.

Newell, D.G., M. Koopmans, L. Verhoef, E. Duizer, A. Aidara-Kane, H. Sprong, and H. Kruse. (2010). Food-borne diseases — The challenges of 20 years ago still persist while new ones continue to emerge. *Int. J.Food Microbiol.* 139:S3-S15.

Noller, A.C., M.C. McEllistrem, A.G.F. Pacheco, D.J. Boxrud, and L.H. Harrison. (2003). Multilocus variable-number tandem repeat analysis distinguishes outbreak and sporadic *Escherichia coli* O157:H7 isolates. *J. Clin. Microbiol.* 41:5389-5397.

Notzon, A., R. Helmuth, and J. Bauer. (2006). Evaluation of an immunomagnetic separationreal-time PCR assay for the rapid detection of *Salmonella* in meat. *J. Food Prot.* 69: 2896-2901.

Nowak, B., T. von Müffling, S. Chaunchom, and J. Hartung. (2007). *Salmonella* contamination in pigs at slaughter and on the farm: A field study using an antibody ELISA test and a PCR technique. *Int. J. Food Microbiol.* 115:259-267.

Nucera, D.M., C.W. Maddox, P. Hoien-Dalen, and R.M. Weigel. (2006) Comparison of API 20E and *invA* PCR for identification of *Salmonella enterica* isolates from swine production units. *J. Clin. Microbiol.* 44:3388–3390.

Ohl, M.E. and S.I. Miller. (2001). Salmonella: A model for bacterial pathogenesis. [Review]. *Annual Review of Medicine.* 52, 259-274.

Okamura, M., Y. Ohba, S. Kikuchi, A. Suzuki, H. Tachizaki, K. Takehara, M. Ikedo, T. Kojima, and M. Nakamura. (2008). Loop mediated isothermal amplification for the rapid, sensitive, and specific detection of the O9 group of *Salmonella* in chickens. *Vet. Microbiol*. 132:197–204.

Olsen, S.J., R. Bishop, F.W. Brenner, T.H. Roels, N. Bean, R.V. Tauxe, and L. Slutsker. (2001). The changing epidemiology of *Salmonella*: Trends in serotypes isolated from humans in the United States, 1987–1997. *J. Infect. Dis.* 183: 753-761.

Olson, M.V. (1989). Separation of large DNA molecules by pulsed-field gel electrophoresis: A review of the basic phenomenology. *J. Chromatogra. A*. 470:377-383.

Parida, M., K. Horioke, H. Ishida, P.K.Dash, P. Saxena, A.M. Jana, M.A. Islam, S. Inoue, N. Hosaka, and K. Morita. (2005). Rapid detection and differentiation of dengue virus serotypes by a real-time reverse transcription-loop-mediated isothermal amplification assay. *J. Clin. Microbiol. 43*:*2895*–2903.

Park, A. (2011). Source of *E. coli* Outbreak-It was the sprouts after all. Available at: http://healthland.time.com/2011/06/10/source-of-e-coli-outbreak-in-germany-%E2%80%94-it-was-the-sprouts-after-all/. Accessed Oct. 21, 2011.

Popoff, M.Y. and L. LeMinor. (1997). Antigenic formulas of the *Salmonella* serovars, 7th revision. World Health Organization Collaborating Centre for Reference and Research on *Salmonella*, Pasteur Institute, Paris, France.

Popoff, M.Y. (2001). Antigenic formulas of the *Salmonella* serovars, 8th revision. W.H.O. Collaborating Centre for Reference and Research on *Salmonella.* Institute Pasteur, Paris, France.

Popoff, M.Y., J. Bockemuhl, and L.L. Gheesling. (2004). Supplement 2002 (no. 46) to the Kauffmann–White scheme. *Res. Microbiol.* 155:568–570.

Pupo, G.M., D.K. Karaolis, R. Lan, and P.R. Reeves. (1997). Evolutionary relationships among pathogenic and nonpathogenic Escherichia coli strains inferred from multilocus enzyme electrophoresis and mdh sequence studies. *Infection and Immunity, 65*(7), 2685-2692.

Rahn, K., S.A. DeGrandis, R.C. Clarke, S.A. McEwen, J.E. Galán, C. Ginocchio, R. Curtiss, and C.L. Gyles. (1992). Amplification of an *invA* gene sequence of *Salmonella* typhimurium by polymerase chain reaction as a specific method of detection of *Salmonella*. *Molec. Cell. Probes.* 6:271-279.

Ramisse, V., P. Houssu, E. Hernandez, F. Denoeud, V. Hilaire, O. Lisanti, F. Ramisse, and G. Vergnaud. (2004). Variable number of tandem repeats in *Salmonella enterica* subsp. *enterica* for typing purposes. *J. Clin. Microbiol.* 42:5722-5730.

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

Reche, M.P., M.A. Echeita, J.E. los Rios, M.A. Usera, P.A. Jimenez, A.M. Rojas, J. Colas, and I. Rodriguez. (2003). Comparison of phenotypic and genotypic markers for characterization of an outbreak of *Salmonella* serotype Havana in captive raptors. *J. Appl. Microbiol*. 94:65–72

Reeves, M.W., G.M. Evins, A.A. Heiba, B.D. Plikaytis, and J.J. Farmer III. (1989). Clonal nature of *Salmonella typhi* and its genetic relatedness to other salmonellae as shown by multilocus enzyme electrophoresis and proposal of *Salmonella bongori* comb. nov. *J. Clin. Microbiol.* 27:313–320.

Rozen, S. and H.J. Skaletsky. (2000). Primer3 on the www for general users and for biological programmes, pp 365-386. *In* S. Krawetz, S. Misener (eds.), Bioinformatics methods and protocols: methods in molecular biology, Humana Press, Totowa.

Rychlík, I., L. Van Kesteren, L. Cardová, A. Švestková, R. Martínková, and F. Šišák. (1999). Rapid detection of *Salmonella* in field samples by nested polymerase chain reaction. *Letters in Applied Microbiology, 29*(4), 269-272.

SantaLucia, J. and D. Hicks. 2004. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* 33:415-440.

Santos, R.L., S. Zhang, R.M. Tsolis, R.A. Kingsley, L.G. Adams, and A.J. Bäumler. (2001). Animal models of *Salmonella* infections: gastroenteritis *vs* typhoid fever. *Microbes Infect.* 3:1335-1344.

Scallan, E., R.M. Hoekstra, F.J. Angulo, R.V. Tauxe, M.-A. Widdowson, S.L. Roy, J.L. Jones, and P.M. Griffin. (2011). Foodborne illness acquired in the United States—major pathogens. *Emerg. Infect. Dis.* 17:7-15

Scheutz, F., E. Møller Nielsen, J. Frimodt-Møller, N. Boisen, S. Morabito, R. Tozzoli, J.P. Nataro, and A. Caprioli. (2011). Characteristics of the enteroaggregative Shiga toxin/verotoxinproducing *Escherichia coli* O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. *Euro. Surveill.* 16:pii=19889. Available online: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19889

Schwartz, D.C. and C.R. Cantor. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell.* 37:67-75.

Scott, F., J. Threlfall, J. Stanley, and C. Arnold. (2001). Fluorescent amplified fragment length polymorphism genotyping of *Salmonella* Enteritidis: a method suitable for rapid outbreak recognition. *Clin. Microbiol. Infect.* 7:479–485.

Smith, C.L. and C.R. Cantor. (1987). Purification, specific fragmentation, and separation of large DNA molecules, pp. 449-467. *In* W. Ray (Ed.), Methods in Enzymology. Academic Press.

Sosin, D.M. (2003). Draft framework for evaluating syndromic surveillance systems. *J. Urban Health,* 80:i8-13.

Soumet, C., G. Ermel, P. Fach and P. Colin. (1994). Evaluation of different DNA extraction procedures for the detection of*Salmonella*. *Lett. Appl. Microbiol.*, 19, pp. 294–298.

Stone, G.G., R.D. Oberst, M.P. Hays, S. McVey, and M.M. Chengappa. (1994). Detection of *Salmonella* serovars from clinical samples by enrichment broth cultivation-PCR procedure. *J. Clin. Microbiol.* 32:1742-1749.

Swaminathan, B. and P. Feng. (1994). Rapid detection of food-borne pathogenic bacteria. *Ann. Rev. Microbiol.*48:401-426.

Swaminathan, B. and T.J. Barrett. (1995) Amplification methods for epidemiologic investigations of infectious disease. *J. Microbiol. Methods.* 2:129–139

Swaminathan, B., T.J. Barrett, S.B. Hunter, and R.V. Tauxe. (2001). PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerg. Infect. Dis.* 7:382–389.

Swaminathan, B., P. Gerner-Smidt, L.-K. Ng, S. Lukinmaa, K.-M. Kam, S. Rolando, and N. Binsztein. (2006). Building PulseNet International: An interconnected system of laboratory networks to facilitate timely public health recognition and response to foodborne disease outbreaks and emerging foodborne diseases. *Foodborne Pathog. Dis. 3*: 36-50.

Tamada, Y., Y. Nakaoka, K. Nishimori, A. Doi, T. Kumaki, N. Uemura, K. Tanaka, S.-I. Makino, T. Sameshima, M. Akiba, M. Nakezawa, and I. Uchida. (2001). Molecular typing and epidemiological study of *Salmonella enterica* serotype Typhimurium isolates from cattle by fluorescent amplified-fragment length polymorphism fingerprinting and pulsed-field gel electrophoresis. *J. Clin. Microbiol*. 39:1057–1066.

Tauxe, R.V. (2002). Emerging foodborne pathogens. *Int. J. Food Microbiol.* 78:31-41.

Tenover, F.C., R.D. Arbeit, R.V. Goering, P.A. Mickelsen, B.E. Murray, D.H. Persing, and B. Swaminathan. (1995). Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J. Clin. Microbiol.* 33:2233–2239.

Thomas, A., T. Cheasty, J.A. Frost, H. Chart, H.R. Smith, and B. Rowe. (1996). Vero cytotoxin-producing *Escherichia coli*, particularly serogroup O157, associated with human infections in England and Wales. *Epidemiol. Infect.* 117:1–10.

Török, T.J., R.V. Tauxe, R.P. Wise, J.R. Livengood, R. Sokolow, S. Mauvais, S., and L.R. Foster. (1997). A large community outbreak of salmonellosis caused by intentional contamination of restaurant salad bars. *JAMA: J. Am. Med. Assoc.* 278:389-395.

United States Department of Agriculture. (2000). National Antimicrobial Resistance Monitoring System for Enteric Bacteria 1999 Final Report. United States Department of Agriculture, Athens, GA.

Urwin, R. and M.C.J. Maiden. (2003). Multi-locus sequence typing: a tool for global epidemiology. *Trends Microbiol.* 11: 479-487.

Van Kessel, J.S., J.S. Karns, and M.L. Perdue. (2003). Using a portable real-time PCR assay to detect *Salmonella* in raw milk. *J. Food Prot.* 66: 1762-1767.

Vázquez-Novelle, M.D., A.J. Pazos, M. Abad, J.L. Sánchez, and M.L. Pérez-Parallé. (2005). Eight-hour PCR-based procedure for the detection of *Salmonella* in raw oysters. *FEMS Microbiol. Lett.* 243: 279-283.

Velusamy, V., K. Arshak, O. Korostynska, K. Oliwa, and C. Adley. (2009). An overview of foodborne pathogen detection: In the perspective of biosensors. *Biotechnol. Adv.* 28:232-254.

Viazis, S. and F. Diez-Gonzalez. (2011). Chapter one - Enterohemorrhagic Escherichia coli: The Twentieth Century's Emerging Foodborne Pathogen: A Review, pp. 1-50. *In* L. S. Donald (ed.), Advances in Agronomy. Academic Press.

Voetsch, A.C., T.J. Van Gilder, F.J. Angulo, M.M. Farley, S. Shallow, R. Marcus, P.R. Cieslak, V.C. Deneen, R.V. Tauxe, and f.t.E.I.P.F.W. Group. (2004). FoodNet estimate of the burden of illness caused by nontyphoidal *Salmonella* infections in the United States. *Clin. Infec. Dis.* 38:S127-S134.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, and M. Kuiper. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23:4407–4414.

Waage, A.S., T. Vardund, V. Lund, and G. Kapperud. (1999). Detection of low numbers of *Salmonella* in environmental water, sewage and food samples by a nested polymerase chain reaction assay. *J. Appl. Microbiol.* 87: 418-428.

Wahlberg, J., J. Lundeberg, T. Hultman, and M. Uhlen. (1990). General colorimetric method for DNA diagnostics allowing direct solid-phase sequencing of the positive samples. *Proc. Natl. Acad. Sci.* 87:6569-73

Wang, L., T. Li, and A. Mustapha. (2007). Rapid and simultaneous quantitation of *Escherichia coli* O157:H7, *Salmonella*, and *Shigella* in ground beef by multiplex real-time PCR and immunomagnetic separation. *J. Food Prot.* 70: 1366-1372.
White, B. (1926). Further studies of the *Salmonella* group. *Med. Res. Counc. Spec. Rep. Ser.* 103:3-160.

Widjojoatmodjo, M.N., A.C. Fluit, R. Torensma, B.H. Keller, and J. Verhoef. (1991). Evaluation of the magnetic immuno PCR assay for rapid detection of *Salmonella*. *Eur. J. Clin. Microbiol. Infect. Dis*. 10:935–938.

Witonski, D., R. Stefanova, A. Ranganathan, G.E. Schutze, K.D. Eisenach, and M.D. Cave, M. D. (2006). Variable number tandem repeats of *Salmonella enterica* subsp.

enterica useful in genotyping isolates of serovars Typhimurium and Newport. *J. Clin. Microbiol.*, *42*:5722-5730.

Wolffs, P.F.G., K. Glencross, R. Thibaudeau, and M.W. Griffiths. (2006) Direct quantitation and detection of Salmonellae in biological samples without enrichment, using two-step filtration and real-time PCR. *App. Environ. Microbiol.* 72:3896–3900.

World Health Organization. (1999). World Health Organization, 1999. Zoonotic non-O157 Shiga toxin-producing *Escherichia coli* (STEC). Report of a WHO Scientific Working Group Meeting. Berlin, Germany, 22–23 June, 1998, WHO, pp 1-30.

World Health Organization. (2007). Food safety & food-borne illness. Fact sheet no. 237. Geneva: World Health Organization.

Wu, C.-J., P.-R. Hsueh, and W.-C. Ko. (2011). A new health threat in Europe: Shiga toxin–producing *Escherichia coli* O104:H4 infections. *J. Microbiol. Immunol. Infect.* 44:390-393.

Wu, D., P. Hugenholtz, K. Mavromatis, R. Pukall, E. Dalin, N.N. Ivanova, Natalia… and J.A. Eisen. (2009). A phylogeny-driven genomic encyclopedia of Bacteria and Archaea. *Nature.* 462:1056-1060.

Zhao, S., D.G. White, and S.L. Friedman. (2006). Characterization of *Salmonella* Heidelberg from retail meat samples: National Antimicrobial Resistance Monitoring System (NARMS): 2002–2004. Abstract from the 106th American Society for Microbiology General Meeting, Orlando, FL.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31:3406–3415.

CHAPTER III

PRIMERS WITH 5' FLAPS IMPROVE THE EFFICIENCY AND SENSITIVITY OF MULTIPLEX PCR ASSAYS FOR THE DETECTION OF *SALMONELLA* AND *ESCHERICHIA COLI* O157:H7

ABSTRACT

Foodborne illnesses caused by *Salmonella enterica* and *Escherichia coli* O157:H7 are worldwide health concerns. Rapid, sensitive, and robust detection of these pathogens in foods and in clinical and environmental samples is essential for routine food quality testing, effective surveillance, and outbreak investigations. The aim of this study was to evaluate the effect on PCR sensitivity of adding a short, AT-rich overhanging nucleotide sequence (flap) to the 5' end of PCR primers specific for the detection of *Salmonella* and *E. coli* O157:H7. Primers targeting the *inv*A gene of *Salmonella* and the *rfbE* gene of *E. coli* O157:H7 were synthesized with or without a 12-bp, AT-rich 5' flap (5'-AATAAATCATAA-3'). Singleplex PCR, multiplex PCR, and real-time PCR sensitivity assays were conducted using purified bacterial genomic DNA and crude cell lysates of bacterial cells. The effect of background flora on detection was evaluated by spiking tomato and jalapeno pepper surface washes with *E. coli* O157:H7 and *Salmonella* Saintpaul. When targeting individual pathogens, end-point PCR assays using flap-

amended primers were more efficient than thos using non-amended primers, with 20.4% and 23.5% increases in amplicon yield for *Salmonella* and *E. coli* O157:H7, respectively. In multiplex PCR assays, a 10- to 100-fold increase in detection sensitivity was observed when the primer flap sequence was incorporated. This improvement in both singleplex and multiplex PCR efficiency and sensitivity can lead to improved *Salmonella* and *E. coli* O157:H7 detection.

INTRODUCTION

Foodborne illnesses cause significant worldwide morbidity and mortality each year (Flint *et al.,* 2005). In the United States, 31 major foodborne pathogens are estimated to cause 9.4 million illnesses, 55,961 hospitalizations, and 1,351 deaths annually (Scallan *et al.,* 2011). Alleviating the significant health and economic burden caused by foodborne pathogens requires effective methods for disease surveillance, outbreak investigations, and contaminated food source identification. Accurate detection of foodborne pathogens is of critical importance but continues to be a challenge. Strategies for detecting human foodborne pathogens should allow detection of most or all strains of a pathogen and should allow accurate identification in a variety of substrates or media. Ideal detection methods should be sensitive, specific, rapid, robust, reproducible, and cost effective, and have the potential for automation (Malorny *et al.,* 2003). Improved detection techniques are needed also for biosecurity and microbial forensics applications, in which rapid and robust detection methods are essential. In the event of

an intentionally-incited disease outbreak, detection is likely to be the first and most critical step in traceback and attribution (Breeze *et al.,* 2005).

Although the International Organization for Standardization (ISO) supports the use of several traditional detection methods for major foodborne pathogens (Anonymous, 1993; Anonymous, 2001), significant limitations exist with many of these culture-based methods (Malorny, 2003). PCR is commonly used for detection of bacterial pathogens in food and clinical samples and has several advantages over traditional methods including decreased detection times, the reliance on genotypic rather than variably expressed phenotypic characteristics, and the ability to detect viable-but-nonculturable cells (Malorny, 2003). Additionally, PCR has become increasingly more popular for human pathogen detection because of its sensitivity, diagnostic accuracy, high detection probability, high robustness, and general acceptance due to user-friendly protocols for application and interpretation (Khan *et al.,* 2011; Malorny, 2003, Malorny *et al.,* 2003; Velusamy *et al.,* 2010). Many techniques for bacterial subtyping also rely on PCR (Wattiau *et al.,* 2011). Improvements in PCR for pathogen detection are primarily in the areas of specificity, sensitivity, and efficiency.

Strains of *Salmonella enterica* subsp. *enterica* and *Escherichia coli* O157:H7, which are among the most important foodborne bacterial pathogens in the United States (Rangel *et al.,* 2005; Voetsch *et al.,* 2004), have both recently been associated with major foodborne outbreaks (CDC, 2011). Rapid identification of foodborne pathogens is essential to minimize further infections, and reducing detection time and increasing sensitivity have been the foci of recent research. The development of multiplex PCR, which allows multiple pathogens to be detected in a single assay, and real-time PCR,

which combines amplification and detection in a single step that can be evaluated in real time, have allowed pathogens to be detected more rapidly and at lower titer. Sometimes even small changes in a PCR procedure can improve the effectiveness of the assay. Afonina *et al.* (2007) found that the addition of short, noncomplementary, AT-rich nucleotide sequences ("flaps") to the 5' end of PCR primers increased the sensitivity of real-time PCR, especially when used with templates difficult to amplify, such as viral and bisulfite-treated DNA (Afonina *et al.,* 2007). Primer flaps were used previously to add restriction sites (Espelund and Jacobsen, 1992) or universal detection sites (Li *et al.,* 2006) to PCR products, and reduce the number of errors when sequencing short PCR products (Binladen *et al.,* 2007). The aim of this study was to evaluate the effect of adding a 5' flap to PCR primers specific for the detection of *Salmonella enterica* and *E. coli* O157:H7 on PCR efficiency and sensitivity. The procedure was assessed in conventional singleplex and multiplex PCR assays, and in a singleplex real-time PCR assay.

MATERIALS AND METHODS

Bacterial strains, culture media, growth conditions. The bacterial strains used in this study were *Salmonella enterica* subspecies *enterica* serovar Typhimurium (cantaloupe isolate), *Salmonella enterica* subspecies *enterica* serovar Saintpaul (produce isolate), *Escherichia coli* O157:H7 K3995 (spinach-associated illness outbreak isolate), and *Escherichia coli* O157:H7 ATCC-43888 (attenuated, without *stx* genes). The bacteria were grown aerobically with 100 RPM shaking at 37°C for 24 h in 5 mL Luria-

Bertani broth (LB broth, Difco, Sparks, MD). Serial dilutions of overnight cultures were prepared in sterile 0.1% peptone (Difco, Sparks, MD) to the desired titers, which were confirmed by dilution plating (in sterile 0.1% peptone) on tryptic soy agar (TSA, Difco, Sparks, MD) and incubating at 37°C for 24 h.

DNA template preparation. Pure genomic DNA was extracted from 1 mL overnight LB broth cultures of approximately 10^9 CFU/mL using the Qiagen Blood and Tissue Kit (Qiagen, Valencia, CA) with an RNase treatment. DNA concentration of the pure genomic DNA was measured with a NanoDrop v. 2000 spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE). *Salmonella* and *E. coli* O157:H7 genomic DNA copy numbers were calculated based on the published genome sizes of *Salmonella* serovar Typhimurium (4.9 Mb plus 94 kb virulence plasmid) (McClelland *et al.,* 2001) and *E. coli* O157:H7 (5.5 Mb plus two plasmids of 90 kb and 6 kb) (Hayashi *et al.,* 2001). Purified genomic DNA was then 10-fold serially diluted from approximately 50 ng/µL (10^9 genomic DNA copies/mL) to approximately 50 ag/µL (10^0 genomic DNA copies /mL). Crude cell lysates of each strain were obtained by boiling 0.5 mL each of 9 10-fold serial dilutions (10^9 to 10^0 CFU/mL) from overnight culture for 10 min. Cellular debris was eliminated at 10,000 rpm for 3 min and supernatant was used for PCR.

Spiked fresh produce background test. To test whether the presence of other niche-sharing bacteria would affect assay results, surface washes from fresh tomatoes and jalapeno peppers were spiked with *Salmonella* and *E. coli* O157:H7. Eighteen tomatoes and 18 jalapeno peppers were locally purchased, weighed, and added three at a time to sterile 400 ml stomacher bags. Fifty ml of Universal Preenrichment Broth (UPB, Difco, Sparks, MD) was added to each of 12 bags (6 bags for each fruit). Each bag was shaken

by hand for 1 min, the surface of each tomato/pepper was then massaged in the bag by hand for 1 min, and finally each bag was again shaken by hand for 1 min. Wash fluid was transferred to sterile containers, serially diluted 10-fold with 0.1% sterile peptone, and plated in duplicate on plate count agar (PCA, Difco, Sparks, MD) for total aerobic counts. Overnight cultures of *Salmonella* and *E. coli* O157:H7 were serially diluted 10-fold in wash fluid from tomato or jalapeno pepper. Crude cell lysates of each spiked wash fluid were obtained by boiling 0.5 mL each of nine 10-fold serial dilutions (10^9 to 10^0 CFU/mL) for 10 min. Cellular debris was eliminated at 10,000 rpm for 3 min.

PCR primers. Commonly used and sensitive PCR primer sets from the published literature were used to evaluate the effects of 5' flap addition. PCR primers targeting the *invA* gene, encoding a protein component of a type three secretion system associated with cellular invasion, were chosen for detection of *Salmonella*. Forward primer, 139 (5'-GTG AAA TTA TCG CCA CGT TCG GGC AA-3'), and reverse primer, 141 (5'-TCA TCG CAC CGT CAA AGG AAC C-3'), amplify a 284 bp fragment (Rahn *et al.,* 1992). PCR primers targeting the *rfbE* gene, encoding an enzyme involved in the synthesis of a component of the O antigen complex specifically binding with the O157 antibody, were chosen for detection of *E. coli* O157:H7 strains. Forward primer, Gi-O157-I (5'-CGA GTA CAT TGG CAT CGT G-3'), and reverse primer, Gi-O157-II (5'-ATT GCG CTG AAG CCT TTG-3'), amplify a 479 bp fragment (Abdulmawjood *et al.,* 2003). All primers were synthesized by Integrated DNA Technologies (Coralville, IA) with and without the 12-bp 5' flap sequence (Table 1), and were validated *in silico,* against available sequences of *Salmonella enterica* subsp. *enterica* and *E. coli* O157:H7, in the GenBank database using BLASTn (Altschul *et al.,* 2007).

PCR amplification. Identical PCR protocols were used to compare the sensitivity and total amplified PCR yield using primers with and without 5' flaps. Twenty µL PCR reaction mixtures contained 1 µL of 10-fold serially diluted template DNA (at concentrations equivalent to 10^8 CFU/mL to 10^0 CFU/mL), 1 µL of each forward and reverse primer (5µM), 10 µL GoTaq Green Master Mix (2X GoTaq Green reaction buffer, 400 µM of each 4 dNTPs, and 3 mM $MgCl_2$) (Promega, Madison, WI), and 7 µL nuclease free water. The PCR program consisted of initial denaturation at 94°C for 3 minutes and 35 cycles of the following: denaturation at 94°C for 60 sec, annealing at 60°C for 60 sec, and extension at 72°C for 60 sec, and followed by a final extension at 72°C for 7 min. PCR amplicons were visualized on 1.5% agarose gels containing SYBR Safe DNA gel stain (Invitrogen, Carlsbad, CA). PCR products amplified using primers with and without 5' flaps were visualized on a single 1.5% agarose gel. Water as a negative control (NTC; non-template control) was included in each PCR assay.

PCR product quantification. The amplicons obtained using PCR primers with and without 5' flaps were eluted from agarose gels using Quantum Prep Freeze 'N Squeeze Spin Columns (Bio-Rad, Hercules, CA). The concentrations of eluted amplicons were determined using a NanoDrop v. 2000 spectrophotometer. Average DNA concentrations from 3 separate singleplex PCR assays were calculated and compared for both *Salmonella* Typhimurium and *E. coli* O157:H7 K3995, with and without flap-amended primers.

Real time PCR. SYBR Green real-time PCR assays were performed using primers with and without 5' flaps in 20 µL reaction mixtures containing 10 µL of Platinum SYBR Green real-time PCR SuperMix-UDG (Invitrogen, Carlsbad, CA), 0.80

μL (5 μM) of each primer, 0.12 μL of bovine serum albumin (BSA, 50 mg/μL) (Invitrogen, Carlsbad, CA), 1 μL of template DNA, and 7.28 μL of nuclease free water. Template DNA used for qPCR was at concentrations equivalent to 10^8, 10^5, and 10^3 CFU/mL for *Salmonella* Typhimurium and *E. coli* O157:H7 K3995. A negative control (water) was included in each PCR assay, and each reaction was performed in three replicates. Cycling parameters included two initial holds, 2 min at 50°C and 2 min at 95°C, followed by 35 cycles at 95°C for 15 sec, followed by 60°C for 60 sec. The real-time PCR assays were performed in a Rotor-Gene 6000 thermocycler (Corbett Research, Sydney, Australia). The threshold fluorescence for calculating the cycling threshold number (C_t) was set at 0.2.

Multiplex PCR. Genomic DNA of *Salmonella* Saintpaul and *E. coli* O157:H7 K3995 was combined and amplified together in an endpoint multiplex PCR assay using primers with and without 5' flaps: the 139-141 primer set for *Salmonella* and the Gi-O157-I- Gi-O157II primer set for *E. coli* O157:H7. Two μL of genomic DNA at a concentration equivalent to 10^9 genomic DNA copies/mL (approximately 50 ng/μL) for each strain was combined and 10-fold serially diluted to 10^0 genomic DNA copies/mL (approximately 50 ag/μL) for sensitivity analysis. PCR component concentrations and conditions remained the same as in singleplex PCR except that the effect of $MgCl_2$ concentration in the multiplex PCR was tested at concentrations of 1.5, 2.5, 3.0, 3.5, 4.0, and 4.5 mM in separate assays with *Salmonella* Saintpaul and *E. coli* O157:H7 K3995 genomic DNA.

Sensitivity analysis. PCR sensitivity was determined based on amplicon detection in conventional PCR assays (presence of band as well as PCR product

quantification by elution) and fluorescence intensity in SYBR Green qPCR assays. Detection was designated as positive in the presence of a band after amplicon electrophoresis in 1.5% agarose gels. Sensitivity was expressed as the lowest DNA concentration equivalent to cellular titer detected in each PCR reaction.

Statistical analysis. Six replicate trials were conducted for each experiment. Data were analyzed using general linear model procedures of the Statistical Analysis software version 9.3 (SAS Institute, Inc., Cary, N.C). Duncan's multiple range test was used to determine the significant difference among the groups ($p<0.05$).

RESULTS AND DISCUSSION

Addition of 5' flaps to PCR primers. The target genes and primers used in this study were selected based on their wide applications and inter-laboratory validation (Abdulmawjood *et al.,* 2003; Malorny *et al.,* 2003; Rahn *et al.,* 1992). The in silico analysis of the primers using BLASTn were evaluated against all sequences in the NCBI database before and after the incorporation of 5' flaps to the primers. No non-specific binding of the primers was predicted with any relevant microorganism sequences in the NCBI database after the incorporation of the 5'flaps.

Although the139-141 primer set is commonly used and well validated, it was observed previously to amplify non-specifically a homologous gene in some *E. coli* strains, and did not allow detection of some *Salmonella* Saintpaul strains (1 out of 20 tested *Salmonella* Saintpaul strains was not detected) (Malorny *et al.,* 2003; Rahn *et al.,* 1992). However, no non-specific fragments or failure to detect *Salmonella* Saintpaul was

observed in this study, whether 5' flap-amended primers were used or not. Additionally, no difference in detection sensitivity was observed between *Salmonella* Saintpaul and Typhimurium when using the 139-141 primer set, or between *E. coli* O157:H7 K3995 and *E. coli* O157:H7 ATCC-43888 when using the Gi-O157-I-Gi-O157-II primer set.

Singleplex PCR. Singleplex endpoint PCR sensitivity with 5' flap-amended primers was not different from PCR assays using primers without flaps when using purified genomic DNA (Figure 1) or 10-fold serially diluted crude cell lysate (data not shown). The limit of detection was 10^2 CFU/mL (ca.5.0 fg/µL) for *Salmonella* and 10^3 CFU/mL (ca. 50 fg/µL) for *E. coli* O157:H7. However, detection was always possible at 10^3 CFU/mL for *Salmonella* (10 copies of genomic DNA per reaction, approximately 50 fg) and 10^4 CFU/mL for *E. coli* (100 copies of genomic DNA per reaction, approximately 500 fg). These results are comparable to previous studies using the 139-141 primer set for *Salmonella* detection (Malorny *et al.,* 2003) and the Gi-O157-I-Gi-O157-II primer set for *E. coli* O157 detection (Abdulmawjood *et al.,* 2003).

The efficiency of the *Salmonella* and *E. coli* O157:H7 singleplex PCR assays using primers with and without flaps was further evaluated by quantifying the concentration of DNA from eluted bands. A 23.5%, and 20.4% increase in yield was observed when 5' flaps were added to primers specific for *E. coli* O157:H7 and *Salmonella*, respectively (Table 2). Although the 12 bp flap sequence is incorporated into the amplified fragment sequence after the first few PCR cycles, this would account for only an approximately 5-10 percent yield increase. Increased primer specificity to template DNA when 5' flaps are used may account for the increase in yield. These results

indicate that amending the 139-141 and Gi-O157-I-Gi-O157-II primer sets used in this study with 5' flap sequences increased the amplification efficiency.

Spiked fresh produce background test. Background flora washed from the surfaces of locally purchased tomatoes and jalapeno peppers had no effect on the detection sensitivity of *Salmonella* Saintpaul and Typhimurium or of *E. coli* O157:H7 K3995 and ATCC-43888. Background bacteria at a concentration of 10^6 CFU/tomato and 10^7 CFU/pepper were present in surface washes before spiking the washes with *Salmonella* or *E. coli* O157:H7 (data not shown).

Real-time PCR. Real-time PCR assays using genomic DNA from *E. coli* O157:H7 K3995 and *Salmonella* Typhimurium, at concentrations equivalent to 10^8, 10^5 and 10^3 CFU/mL, and primers with and without 5' flaps showed an increase in fluorescence intensity when 5' flaps were used (Figure 2). However, the C_t values between amplification with 5' flap amended and non-amended primers at each concentration were not significantly different ($p > 0.01$, data not shown). The increased fluorescent signal with 5' flap amended primers may be a combined result of the 24 bp (from the flap) larger amplicon size and the increased PCR amplicon yields shown above.

Multiplex PCR. Initial multiplex PCR assays using GoTaq Green master mix (containing 1.5 mM $MgCl_2$) yielded little amplification of *E. coli* O157:H7 at lower cell concentration, with or without 5' flap-amended primers. Subsequent optimization of the $MgCl_2$ concentration to 3.0 mM for multiplex PCR resulted in increased assay sensitivity for both pathogens but particularly for *E. coli* O157:H7 with 5' flaps (Figure 3). Optimization of the annealing temperature in multiplex was not attempted since extensive optimization studies for the *rfbE* primer set (Abdulmawjood *et al.,* 2003) and the *invA*

primer set (Rahn *et al.,* 1992) have previously been done. There was a 10-and 100-fold increase in the detection limit for *E. coli* O157:H7 and *Salmonella*, respectively, in multiplex PCR with 5' flaps (Figure 4A). The detection limit of multiplex PCR with 5' flaps also increased for each pathogen compared to singleplex PCR. When only one primer set was used with the mixture of *Salmonella* and *E. coli* O157:H7 genomic DNA, no non-specific binding was observed and the level of detection observed was the same for each pathogen (Figure 4B). The reason for the 10- to 100-fold increase in multiplex PCR sensitivity is unknown, but was observed repeatedly. The limit of detection observed in multiplex was 10^2 CFU/mL for *E. coli* O157:H7 and 10^1 CFU/mL for *Salmonella* (Figure 4). These concentrations correspond to 1 and 0.1 genomic copies per PCR reaction for *E. coli* and *Salmonella*, respectively (assuming 100% recovery of purified genomic DNA). These results may represent the highest level of sensitivity possible for the detection of the foodborne pathogens *Salmonella* and *E. coli* O157:H7.

Increased $MgCl_2$ concentration in singleplex PCR from 1.5 mM to 3.0 mM resulted in no difference in PCR sensitivity and detection probability (data not shown). Increased assay sensitivity with optimized $MgCl_2$ concentration was observed only in multiplex PCR. Mg^{+2} ions form complexes with dNTPs in PCR reactions, allowing proper elongation by Taq polymerase, and promote DNA/DNA interactions. The increase in sensitivity may be a result of increased efficiency of primer to target annealing when Mg^{+2} concentrations are optimal.

In summary, the addition of a 12 bp AT-rich 5' flap sequence to PCR primers specific for *Salmonella* and *E. coli* O157:H7 increased the detection sensitivity 10- to 100-fold in multiplex PCR compared to that of assays done using primers without flaps.

In singleplex PCR, 5' flap addition to PCR primers led to no significant difference in sensitivity but allowed higher amplicon yield using the selected primer sets. The results suggest that adding AT-rich 5' flap sequences to PCR primers may improve detection of pathogens by multiplex PCR. The method has potential applications in the food and agricultural industries, foodborne outbreak surveillance, pathogen identification and traceback, and biosecurity and microbial forensics.

LITERATURE CITED

Abdulmawjood, A., M. Bülte, N. Cook, S. Roth, H. Schönenbrücher, and J. Hoorfar. (2003). Toward an international standard for PCR-based detection of *Escherichia coli* O157: Part 1. Assay development and multi-center validation. *J. Microbiol. Methods.* 55: 775-786.

Afonina, I., I. Ankoudinova, A. Mills, S. Lokhov, P. Huynh, and W. Mahoney. (2007). Primers with 5′ flaps improve real-time PCR. *BioTechniques.* 43:770-774.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. (1990). Basic local alignment search tool. *J. Mol. Biol*. 215:403–410.

Anonymous. (1993). Microbiology of food and animal feeding stuffs—horizontal method for the detection of *Salmonella* (ISO 6579:1993). International Organization for Standardization, Geneva, Switzerland.

Anonymous. (2001). Microbiology of food and animal feeding stuffs—horizontal method for the detection of *Escherichia coli* O157 (ISO 16654:2001). International Organization for Standardization, Geneva, Switzerland.

Binladen, J., M.T. Gilbert, P.F. Campos, and E. Willerslev. (2007). 5′-tailed sequencing primers improve sequencing quality of PCR products. *BioTechniques.* 42:174-176.

Breeze, R.G., B. Budowle, and S.E. Schutzer (eds.). (2005). Microbial forensics. Academic Press, San Diego, CA.

Centers for Disease Control and Prevention. (2011). Multistate foodborne outbreaks. Available at: http://www.cdc.gov/outbreaknet/outbreaks.html. Accessed 24 August 2011.

Espelund, M. and K.S. Jacobsen. (1992). Cloning and direct sequencing of plant promoters using primer-adapter mediated PCR on DNA coupled to a magnetic solid phase. *BioTechniques.* 13:74-81.

Flint, J.A., Y.T. Van Duynhoven, F.J. Angulo, S.M. DeLong, P. Braun, M. Kirk, E. Scallan, M. Fitzgerald, G.K. Adak, P. Sockett, A. Ellis, G. Hall, N. Gargouri, H. Walke and P. Braam. (2005). Estimating the burden of acute gastroenteritis, foodborne disease, and pathogens commonly transmitted by food: an international review. *Clin. Infec. Dis.* 41:698-704.

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

Khan, J. A., R.S. Rathore, I. Ahmad, and S. Khan. (2011). Molecular strategies: detection of foodborne bacterial pathogens, p 189-206. *In* I. Ahmad, F. Ahmad and J. Pichtel (eds.). Microbes and Microbial Technology. Springer, New York.

Li, X., Y. Huang, Y. Guan, M. Zhao, and Y. Li. (2006). Universal molecular beacon-based tracer system for real-time polymerase chain reaction. *Anal. Chem*. 78:7886-7890.

McClelland, M., K.E. Sanderson, J. Spieth, S.W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston, and R.K. Wilson. (2001). Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature.* 413:852–856.

Malorny, B. (2003). Standardization of diagnostic PCR for the detection of foodborne pathogens. *Int. J. Food Microbiol.* 83:39-48.

Malorny, B., J. Hoorfar, C. Bunge, and R. Helmuth. (2003). Multicenter validation of the analytical accuracy of *Salmonella* PCR: towards an international standard. *Appl. Environ. Microbiol.* 69:290-296.

Malorny, B., E. Paccassoni, P. Fach, C. Bunge, A. Martin, and R. Helmuth. (2004). Diagnostic real-time PCR for detection of *Salmonella* in food. *Appl. Environ. Microbiol.* 70:7046-7052.

Rahn, K., S.A. DeGrandis, R.C. Clarke, S.A. McEwen, J.E. Galán, C. Ginocchio, R. Curtiss, and C.L. Gyles. (1992). Amplification of an *invA* gene sequence of *Salmonella* Typhimurium by polymerase chain reaction as a specific method of detection of *Salmonella*. *Mol. Cell. Probes.* 6:271-279.

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

SantaLucia, J. and D. Hicks. (2004). The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* 33:415-440.

Scallan, E., R.M. Hoekstra, F.J. Angulo, R.V. Tauxe, M.-A. Widdowson, S.L. Roy, J.L. Jones, and P.M. Griffin. (2011). Foodborne illness acquired in the United States—major pathogens. *Emerg. Infect. Dis.* 17:7-15

Velusamy, V., K. Arshak, O. Korostynska, K. Oliwa, and C. Adley. (2010). An overview of foodborne pathogen detection: In the perspective of biosensors. *Biotechnol. Adv.* 28:232-254.

Voetsch, A.C., T.J. Van Gilder, F.J. Angulo, M.M. Farley, S. Shallow, R. Marcus, P.R. Cieslak, V.C. Deneen, R.V. Tauxe, and f.t.E.I.P.F.W. Group. (2004). FoodNet estimate of the burden of illness caused by nontyphoidal *Salmonella* infections in the United States. *Clin. Infec. Dis.* 38(Supplement 3):S127-S134.

Wattiau, P., C. Boland, and S. Bertrand. (2011). Methodologies for *Salmonella enterica* subsp. *enterica* subtyping: Gold standards and alternatives. *Appl. Environ. Microbiol.* 77:7877-7885.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31:3406–3415.

TABLES

Table 1. PCR primers, with and without 5' flap sequences, targeting the *inv*A gene of *Salmonella* and the *rfb*E gene of *E. coli* O157:H7 strains, used in this study.

Primer name	Target gene	Sequence (5'—3')	Amplicon size (bp)	Reference
139	*inv*A	GTGAAATTATCGCCACGTTCGGGCAA	284	Rahn *et al.,* 1992
141		TCATCGCACCGTCAAAGGAACC		
139-F		[a]**AATAAATCATAA** GTGAAATTATCGCCACGTTCGGGCAA	308	
141-F		[a]**AATAAATCATAA** TCATCGCACCGTCAAAGGAACC		
Gi-O157-I	*rfb*E	CGAGTACATTGGCATCGTG	479	Abdulmawjood *et al*., 2001
Gi-O157-II		ATTGCGCTGAAGCCTTTG		
Gi-O157-I-F		[a]**AATAAATCATAA** CGAGTACATTGGCATCGTG	503	
Gi-O157-II-F		[a]**AATAAATCATAA** ATTGCGCTGAAGCCTTTG		

[a] Bold font indicates 5' flap sequence.

Table 2. Concentration of amplicon DNA in agarose gel bands eluted after singleplex PCR using primers with and without 5' flap sequences, quantified by a NanoDrop spectrophotometer.

Pathogen	Primers	Avg. [DNA] (ng/µl)	Avg. Eluted Vol. (µl)	Avg. [DNA] (ng/band)	Standard Deviation	Difference	p value	% increase in yield
E. coli O157:H7	*rfbE*	9.27	98.07	897.03	53.86	210.4	0.011	23.5
	rfbE-Flap	9.47	117.4	1107.37	61.50			
Salmonella Typhimurium	*invA*	12.40	116.37	1403.60	84.25	286.0	0.010	20.4
	invA-Flap	14.4	118.73	1689.57	69.14			

FIGURES

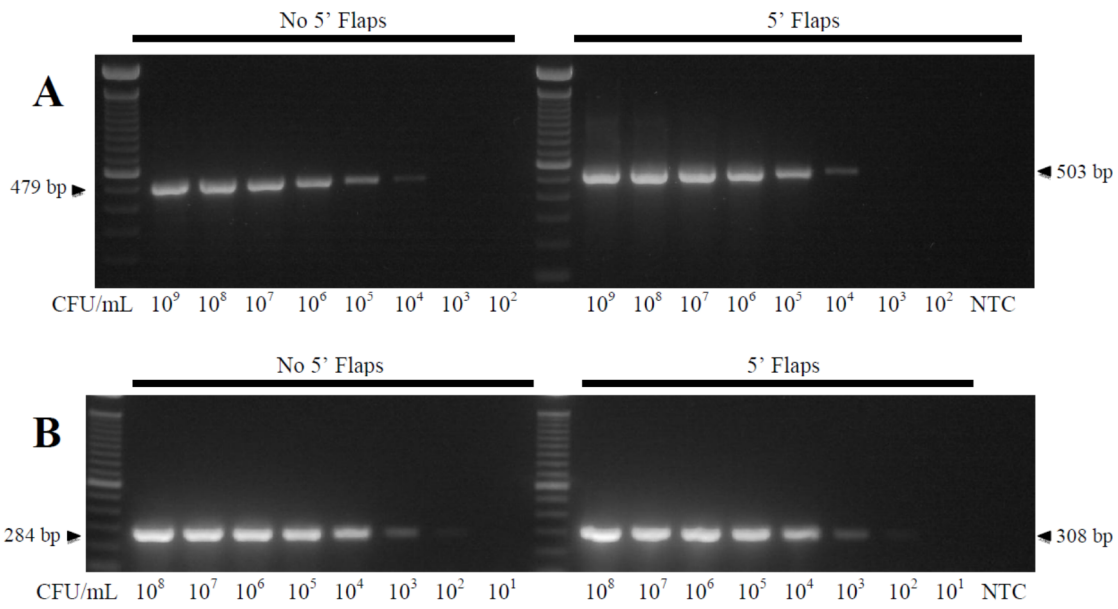

Figure 1. Singleplex PCR detection of *E. coli* O157:H7 K3995 (A) and *Salmonella* Typhimurium (B) 10-fold serially diluted purified genomic DNA, amplified with and without 5' flap amended primers. Lane NTC: no template control.

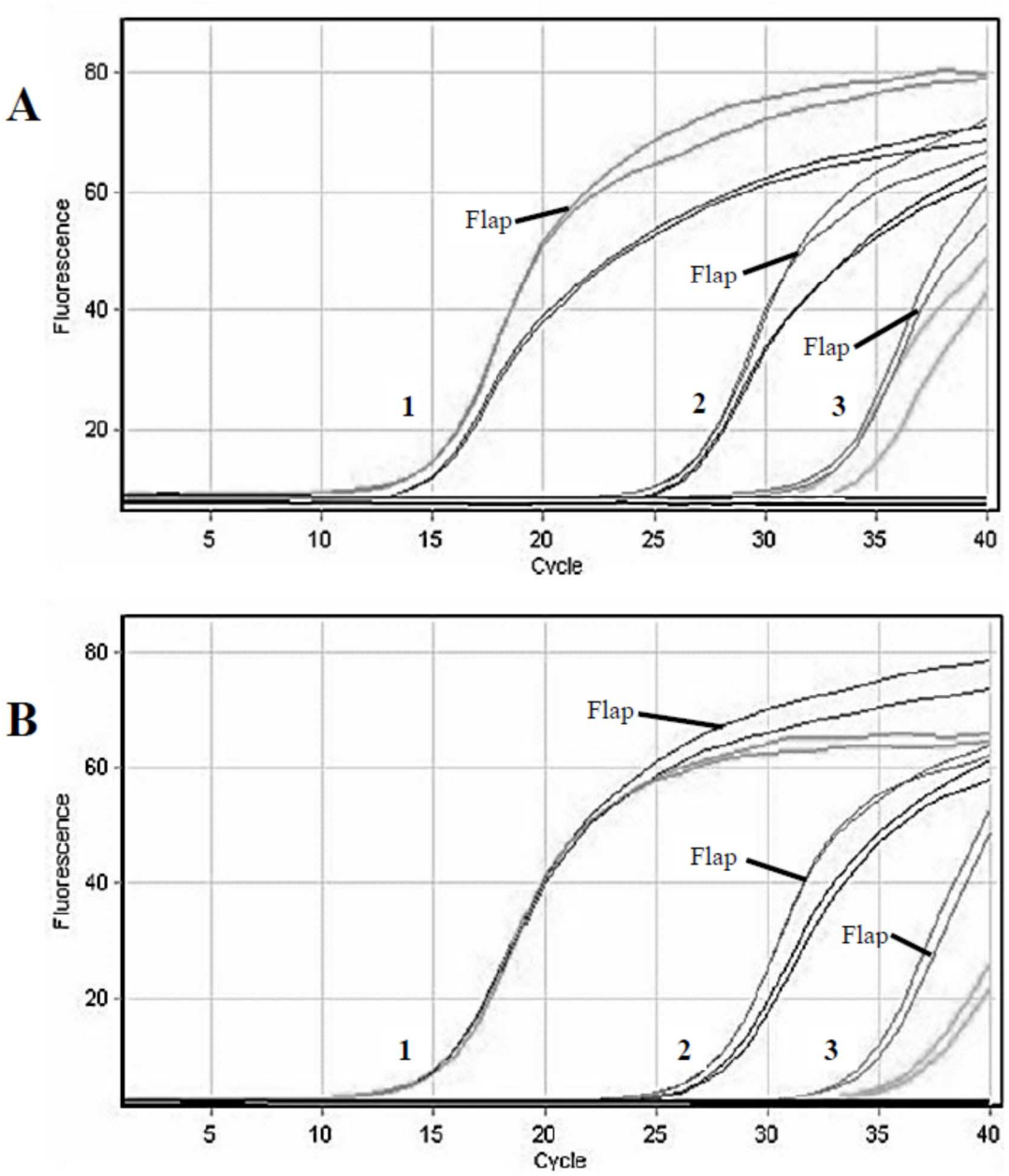

Figure 2. Real-time PCR amplification of genomic DNA of *Salmonella* Typhimurium (A) and *E. coli* O157:H7 K3995 (B). Template DNA, isolated from cultures at cellular concentrations of 10^8 CFU/ml (1), 10^5 CFU/ml (2), and 10^3CFU/ml (3), was amplified in duplicate, with and without 5' flap amended primers.

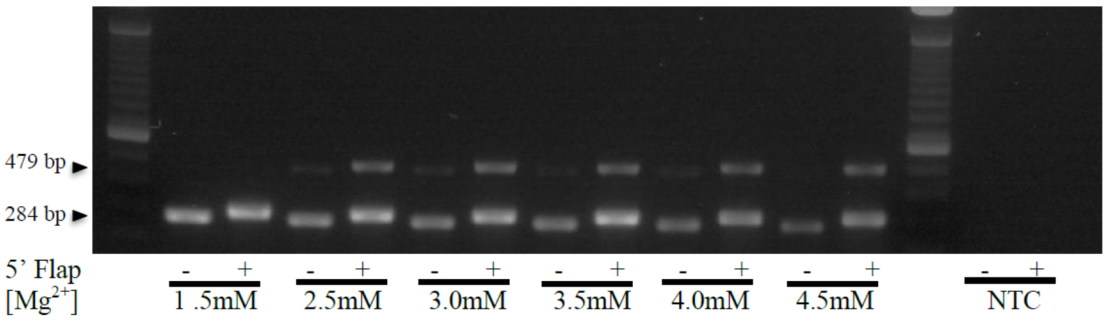

Figure 3. $MgCl_2$ concentration titration using *Salmonella* Saintpaul (284 bp) and *E. coli* O157:H7 K3995 (479 bp) in multiplex using primers with and without 5' flaps. $MgCl_2$ concentration is marked under the lanes. Lane NTC: no template control.

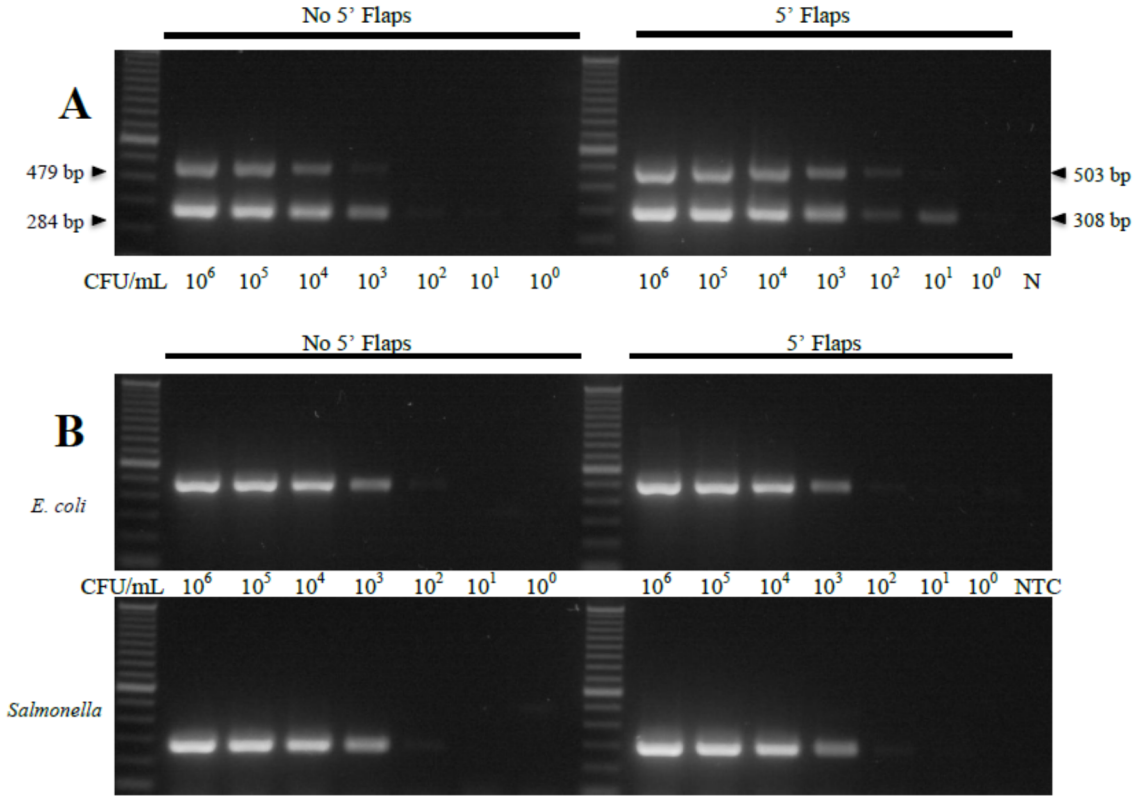

Figure 4. (A) Multiplex detection of 10-fold serially diluted *Salmonella* Saintpaul and *E. coli* O157:H7 K3995 genomic DNA using PCR reaction mixtures with 3.0 mM $MgCl_2$, and primers with and without 5' flaps. (B) Detection of *E. coli* using *rfb*E primers, with and without flaps, in a serially diluted mixture of *E. coli* O157:H7 K3995 and *Salmonella* Saintpaul. (C) Detection of *Salmonella* using *inv*A primers, with and without flaps, in a serially diluted mixture of *E. coli* O157:H7 K3995 and *Salmonella* Saintpaul. Lane NTC: no template control.

CHAPTER IV

MULTIPLE-LOCUS VARIABLE-NUMBER TANDEM REPEAT ANALYSIS FOR STRAIN DISCRIMINATION OF NON-O157 SHIGA TOXIN-PRODUCING *ESCHERICHIA COLI*

ABSTRACT

Non-O157 Shiga toxin-producing *Escherichia coli* (STEC) are emerging pathogens of growing worldwide concern. Recent multistate and multinational outbreaks of foodborne illness have been attributed to consumption of non-O157 STEC-contaminated food. Rapid and sensitive molecular based strain discrimination methods are critical for quick identification of outbreaks and fast traceback to the contaminated food source, hence controlling the number of future infections. The objective of this study was to develop a multiple-locus variable-number tandem repeat (VNTR) analysis (MLVA) assay for intra- and inter-serogroup discrimination of 6 major non-O157 STEC serogroups: O26, O111, O103, O121, O45, and O145. The developed MLVA method consists of 12 VNTR loci in 3 multiplex PCR reactions. Sixty five unique MLVA types were obtained among 84 clinical non-O157 STEC isolates comprised of geographically diverse sporadic and outbreak related strains. Four of the 6 serogroups clustered separately in a minimum spanning tree. The developed MLVA method allowed a higher level of discrimination among serogroups O26, O111, O103, and O121, with discriminatory powers of 1.0, 0.96,

0.97, and 0.97, respectively, than it did for serogroups O45 and O145, with discriminatory powers of 0.90 and 0.82, respectively. Compared to pulsed-field gel electrophoresis (PFGE), a higher level of discrimination was possible for serogroup O26, a similar level of discrimination was observed for serogroups O111, O103, and O121, and a lower level of discrimination was observed for serogroups O45 and O145. The developed non-O157 STEC MLVA method for discrimination of epidemiologically related, highly clonal non-O157 STEC isolates needs to be further validated with more outbreak related isolates and compared to PFGE.

INTRODUCTION

Escherichia coli is a very diverse enteric bacterial species that is an important constituent of the natural gut micro flora of warm-blooded organisms. Most *E. coli* strains are harmless (commensal), but some are pathogenic to humans. Based on the pathogenic mechanisms, at least 5 different pathogenic types of virulent *E. coli* strains exist: Enterohemmorhagic *E. coli* (EHEC), Enteroaggregative *E. coli* (EAEC), Enteroinvasive *E. coli* (EIEC), Enterotoxigenic *E. coli* (ETEC), and Enteropathogenic *E. coli* (EPEC) (CDC, 2012; Todar, 2012; Nataro and Kaper, 1998). The most severe human illness, caused by EHEC, is associated with the production of one or more Shiga toxins (encoded by the genes stx_1 and/or stx_2) and a few other virulence determinants encoded within the locus of enterocyte effacement (LEE) (O'Brien *et al.,* 1992; Ethelberg *et al.,* 2004, Gyles, 2007). Symptoms of Shiga toxin-producing *E. coli* (STEC) infections include severe gastroenteritis, hemorrhagic colitis, or the life-threatening hemolytic

uremic syndrome (HUS) (Besser *et al.,* 1999; Tarr *et al.,* 2005). Of over 100 STEC serogroups associated with human illness by the World Health Organization, STEC O157 is the most commonly isolated in the United States and causes the highest percentage of illnesses (WHO 2012, Scallan *et al.,* 2011; Johnson *et al.,* 1996; CDC, 2012). However, an increased prevalence of non-O157 STEC-induced illnesses has been observed in recent years and non-O157 STEC serogroups may be considered emerging pathogens (Brooks *et al.,* 2005; Johnson *et al.,* 2006; Bettelheim, 2007). Between 2000 and 2005, the global occurrence of non-O157 STEC serogroups causing human illness increased by 60.5% (Coombes *et al.,* 2008). In the United States, the prevalence of non-O157 STEC infections increased steadily from 2000 to 2007 but have remained fairly constant in recent years, while STEC O157 isolation rates are at their lowest since 1998 (CDC, 2012b).

The most commonly isolated non-O157 STEC serogroups associated with human illness vary by year and location (Brooks *et al.,* 2005; CDC, 2012b; CDC, 2012c; Bettelheim, 2007), but serogroups O26, O111, O103, O121, O45, and O145 are usually at the top of the list and are often referred to as the 'big 6' non-O157 STEC serogroups (Karmali *et al.,* 2003). In September of 2011, the United States Department of Agriculture (USDA) Food Safety and Inspection Service (FSIS) identified the 'big 6' non-O157 STEC serogroups as official meat adulterants and the requirement for routine verification tests for these pathogens in raw beef manufacturing trimmings was implemented in June of 2012 (77 *Fed. Reg.* 105, 31 May 2012). Non-O157 STEC serogroups have been isolated from cattle in multiple studies (Wells *et al.,* 1991; Cray *et al.,* 1996; Renter *et al.,* 2005; Cho *et al.,* 2006; Rhoades *et al.,* 2009) and ruminants are

believed to be the primary STEC reservoirs. In addition to consumption of contaminated meat products, outbreaks of foodborne illness caused by non-O157 STEC serogroups have been associated with milk, cheese, water, and fresh produce (Kaspar *et al.,* 2010).

As with most other bacterial foodborne pathogens, sporadic cases of non-O157 STEC infections greatly outnumber outbreak-related cases (Hundley and Cameron, 2004, McPherson *et al.,* 2008; Nielsen *et al.,* 2006; Rivas *et al.,* 2008). Molecular subtyping methods are essential to outbreak investigations from the initial identification of an outbreak to traceback of contamination sources. The PulseNet network coordinated by the Centers for Disease Control and Prevention (CDC) is a national and international molecular subtyping network that functions as a foodborne illness cluster detection tool. Epidemiological investigation of illness clusters allows identification of foodborne illness outbreaks (Swaminathan *et al.,* 2006). The primary molecular subtyping method used by PulseNet is pulsed-field gel electrophoresis (PFGE), commonly referred to as the gold standard subtyping method for pathogenic foodborne bacteria (Swaminathan, 2001). Although the subtyping usefulness and bacterial strain discriminatory capability of PFGE is well documented and demonstrated by the success of the PulseNet network, the technique has several drawbacks. PFGE is a time-consuming and laborious method requiring a high level of technical skill and rigorous standardization to allow inter-laboratory data sharing. In some cases PFGE does not always allow optimal discrimination among closely related bacterial isolates (Hyytia-Trees *et al.,* 2006). To overcome these limitations, PulseNet has begun to augment PFGE data of outbreak-related bacterial isolates with DNA sequence-based methods.

Multiple-locus variable-number tandem repeat (VNTR) analysis (MLVA) is a molecular subtyping method based on differing numbers of tandem repeats within several VNTR loci throughout a bacterial genome (Keim *et al.,* 2000). Following PCR amplification of VNTR loci, the amplified DNA fragments are sized or sequenced and compared among different strains. The size and tandem repeat copy number of each VNTR locus can be designated as a discrete allele number, allowing data comparison among multiple laboratories over extended periods of time (Hyytia-Trees *et al.,* 2006). MLVA is currently used by PulseNet to help discriminate among highly clonal isolates of *Salmonella* Typhimurium (Lindstedt *et al.,* 2003; Lindstedt *et al.,* 2004), *Salmonella* Enteritidis (Cho *et al.,* 2007; Boxrud *et al.,* 2007), and STEC O157 (Hyytia-Trees *et al.,* 2006), with MLVA protocols for *Listeria monocytogenes* and *Salmonella* Newport in active development and validation.

The current STEC O157 MLVA protocol used by PulseNet (Hyytia-Trees *et al.,* 2006), an optimized and modified 9-locus version of the MLVA method developed by Keys e*t al.* (2005), has proven to be useful, allowing a high level of discrimination in conjunction with PFGE. However, this protocol was developed specifically for STEC O157 (especially STEC O157:H7) and does not provide the discriminatory power necessary to discriminate among non-O157 STEC serogroups (Izumiya *et al.,* 2010; Lindstedt *et al.,* 2007). Most previously developed MLVA methods target a single serogroup or serotype and development of a MLVA method for multiple serogroups poses notable challenges (Karma and Gyles, 2010). Discriminatory power at the serotype level is likely to be decreased if multiple serogroups are targeted in a single protocol.

A MLVA genotyping assay that discriminates among many *E. coli* serogroups (not just STEC) was recently developed by Lindstedt *et al.* (2007). The method was used to type all members of the *E. coli* reference (ECOR) collection and 61 human pathogenic *E. coli* and *Shigella* isolates. This novel adaptation of MLVA allowed similar or better discriminatory capability than PFGE with the added benefits of increased assay speed and greater automation. This MLVA method was used to help identify the sources and epidemiological characteristics of an outbreak of STEC O103:H25 (Schimmer *et al.,* 2008) and an outbreak of O145:H28 (Wahl *et al.*, 2011), both in Norway. It was used also by Bustamante *et al.* (2010) and Franci *et al.* (2011) to assess the ability of the method to type and genetically characterize non-O157 STEC serogroups isolated from food, humans, cattle, and beef products in Argentina with a high level of inter-serogroup discrimination. Another MLVA method designed specifically for subtyping of STEC O157, O111, and O26 was developed by Izumiya *et al.* (2010) by essentially adding 9 VNTR loci to the MLVA protocol developed by Hyytia-Trees *et al.* (2006). In both MLVA methods, the complete genomic sequences of one or more STEC O157:H7 serotypes were used to identify VNTR loci, with (Izumiya *et al.,* 2010) or without (Lindstedt *et al.,* 2007) additional comparison with the completed genome sequences of 2 non-O157 STEC serogroups. However, STEC O157 has been shown to form a distinct monophyletic clade among most other *E. coli* serogroups (Reid, 2010) and although a large number of virulence genes are shared, O157 and non-O157 STEC serogroups have independently acquired large numbers of serotype-specific genes by lateral gene transfer (Ogura *et al.,* 2007). It is likely that non-O157 STEC serogroups are evolutionarily more similar to one another than they are to STEC O157 (Donnenberg and Wittam, 2001). To

our knowledge, no MLVA assay has yet been developed specifically for non-O157 STEC serogroups.

The objective of this study was to develop a robust and highly discriminatory MLVA assay for the 6 major non-O157 STEC serogroups—O26, O111, O103, O121, O45, and O145. The concordance of the MLVA data with PFGE data is presented and the MLVA assay is also used to type STEC O157, generic *E. coli*, and EPEC for comparison. The PulseNet MLVA protocol for STEC O157 developed by Hyytia-Trees *et al.* (2006) is one of the most robust and extensively validated MLVA assay for subtyping pathogenic *E. coli*. The technical aspects and details of the PulseNet MLVA protocol for STEC O157 were followed as closely as possible to allow streamlined modification and adaptation of this protocol should it prove useful for molecular subtyping of epidemiologically significant non-O157 STEC serogroups and be adopted by PulseNet.

MATERIALS AND METHODS

Bacterial strains. A total of 92 *E. coli* strains were used in this study. Initial assay development and optimization was done with 24 non-O157 STEC strains obtained from the STEC Center of Michigan State University as part of a non-O157 STEC reference set. This set includes 4 individual strains of each of the 6 major non-O157 STEC serogroups (O26, O103, O111, O121, O145, and O45) isolated from humans (Table 1). Further validation was carried out with 60 non-O157 STEC isolates obtained from CDC (10 strains from each of the big 6 non-O157 serogroups). All 60 strains were

clinical isolates associated with outbreaks or sporadic cases (Table 2). Epidemiological information for all 60 strains and PFGE data for 34 strains were provided by CDC. In addition to the 84 non-O157 STEC isolates, 5 isolates of STEC O157:H7, 2 isolates of EPEC (O55 and O119), and 1 strain of generic *E. coli* K-12 were also analyzed for comparison (Table 3).

VNTR locus selection. To identify potentially useful VNTR loci for inter- and intra-serogroup discrimination of the non-O157 STECs, the published genomes of non-O157 STEC strains O26 (GenBank accession number NC_013361.1), O103 (GenBank accession number NC_013353.1), and O111 (GenBank accession number NC_013364.1) were scanned for tandem repeats using the Tandem Repeats Finder software (Benson, 1999). Selection of a VNTR locus was based on several criteria: a locus had to be present in at least two of the three strains, had to have a high number of tandem repeat percent matches (>80 %), and had to have a low percentage of indels (<3 %). Following initial selection of possible loci, the flanking sequences of each of the VNTR loci were aligned with ClustalW. Only VNTR loci having highly similar flanking sequences were selected to allow optimal primer design and minimize the need for degenerate primers. Additionally, VNTR loci exhibiting differences in tandem repeat copy numbers among the 3 strains were preferentially selected. The more diverse loci (larger difference in copy number) were selected to help discriminate closely related strains within serogroups while the less diverse loci (smaller difference in copy number) were selected to help discriminate among serogroups (Keys *et al.,* 2005). The final selection included 12 VNTR loci, designated SVL (STEC VNTR locus)-1 – SVL-12 (Table 4). One locus,

SVL-11, which was plasmid-located and present only in O111, was selected to evaluate the usefulness of plasmid-located VNTR loci for inter-serogroup discrimination.

The presence and diversity of the selected loci in STEC O157:H7 strains were evaluated also by comparing each selected VNTR locus with the Tandem Repeats Finder results of four published STEC O157:H7 genomes (*E. coli* EDL933 (NC_002655.2), *E. coli* Sakai (NC_002695.1), *E. coli* EC4115 (NC_011353.1), and *E. coli* TW14359 (NC_013008.1)). All loci except SVL-10 and SVL-12 were present also in STEC O157:H7. Locus SVL-1 was previously described by Keys *et al.* (2005) (named O157-2), and SVL-4 by Lindstedt *et al.* (2007) (named CVN004), and SVL-3 was described first by Keys *et al.* (2005) (named O157-11) and later by Lindstedt *et al.* (2007) (named CVN014).

DNA preparation. Bacterial strains were grown overnight at 37°C on trypticase soy agar (TSA). Two to three colonies were suspended in 100 µL of sterile distilled water and boiled for 10 min at 100°C. The suspension was cooled briefly and centrifuged at 10,000 rpm (8165 x *g*) for 10 min. The undiluted supernatant was used as template DNA for PCR amplification and stored at -20°C.

Primer design and PCR amplification. PCR primers for amplification of selected VNTR loci were designed using Primer3 software (Rozen and Skaletsky, 2000) followed by an evaluation of primer thermodynamics using the Mfold web server (Zucker, 2003), and a BLAST search against the NCBI database for primer specificity analysis. PCR primers were designed to minimize multiplex reactions and to allow all multiplex PCR reactions to occur at the same cycling conditions. Therefore, all primers were designed with minimal 3' self-complementary sequences and with similar lengths,

GC contents and melting temperatures. Although loci SVL-1, SVL-3, and SVL-4 were identified previously and used in MLVA protocols for STEC O157:H7, the primers were redesigned to have characteristics similar to those of all other primers in this study. Additionally, MultiPLX 2.1 (Kaplinski *et al.,* 2005) was used to evaluate the potential for primer dimer formation among all 12 primer sets. Since the size range of the amplified fragments for each VNTR locus were unknown, all primers were designed to allow multiplexing of any combination of primer sets. Initial screening of the amplification effectiveness of the 12 primer sets was carried out with the 24-isolate non-O157 STEC reference set and visualized by agarose gel electrophoresis. Based on the amplicon sizes, the primer sets were combined into three multiplex PCR reactions (R1, R2, and R3). Reaction 1 contained primer sets SVL-1, SVL-3, SVL-4, and SVL-8, reaction 2 contained primer sets SVL-2, SVL-6, SVL-10, and SVL-12, and reaction 3 contained primer sets SVL-5, SVL-7, SVL-9, and SVL-11. Forward PCR primers were then fluorescently labeled to allow accurate sizing by multicolor capillary electrophoresis (Table 4).

Unlabeled forward and reverse primers were synthesized by Integrated DNA Technologies (Coralville, IA) and fluorescently labeled forward primers were synthesized by Applied Biosystems (Foster City, CA). The PCR amplification conditions were designed to mimic, as closely as possible, the PCR reaction conditions and reagent concentrations currently used for MLVA by the PulseNet network of the CDC (Hyytia-Trees *et al.,* 2006; Lindstedt *et al.,* 2004; Boxrud *et al.,* 2007). PCR amplification was performed in final volumes of 10 µL consisting of 1.5 µL of 5X Colorless GoTaq Reaction Buffer with 1.5 mM $MgCl_2$ at the 1X concentration (Promega, Madison, WI),

0.4 µL of 50mM $MgCl_2$ (bringing final $MgCl_2$ concentration to 2.0 mM), 1.0 U of GoTaq DNA Polymerase (Promega), 0.2 mM of PCR Nucleotide Mix (Promega), and 1.0 µL of DNA template. Primer concentrations were adjusted to allow optimal peak heights for confident size calling. The amplification conditions consisted of an initial denaturation step at 95°C for 5 min, followed by 35 cycles of 95°C for 30 sec, 58°C for 45 sec, and 72°C for 45 sec, with a final extension step at 72°C for 5 min.

Fragment analysis. Amplified PCR products were diluted 1:60 in sterile distilled water. A 1.0 µL aliquot of the diluted PCR product was added to 9 µL of Hi-Di Formamide (Applied Biosystems) and 0.4 µL of GeneScan 600LIZ size standard (Applied Biosystems). PCR products were sized using an ABI 3730 Genetic Analyzer (Applied Biosystems).

Pulsed-field gel electrophoresis. PFGE was performed by PulseNet according to the standardized PulseNet protocol (Ribot *et al.,* 2006). Among the 60 non-O157 STEC isolates provided by CDC, 36 were analyzed using *Xba*I restriction enzyme (Roche Applied Science, Indianapolis, IN) and 18 were analyzed using *Bln*I restriction enzyme (Roche Applied Science). PFGE patterns were analyzed with BioNumerics software version 4.0 (Applied Maths, Kortrijk, Belgium), uploaded to the PulseNet PFGE pattern database, and named according to the standard nomenclature system (Swaminathan, 2001). The names of the PFGE patterns were provided with the isolates by CDC.

Data analysis. Fragment data were evaluated with GeneMapper software (Applied Biosystems) and fragment size lists were imported to BioNumerics software version 6.6 (Applied Maths) for analysis. BioNumerics links fragment data to corresponding isolate information previously imported. The tandem repeat (TR) copy

number of each locus is calculated based on the fragment size in comparison to the offset size. Partial repeats were rounded up or down to the closest complete TR number in accordance with Hyytia-Trees *et al.* (2006). For each locus, alleles were named according to the number of TRs, whereas null alleles, as no PCR amplification at a given locus, are designated as -2.0 to differentiate between null alleles and VNTR loci with no TRs. The diversity index (D_N) for each locus was calculated based on Simpson's diversity index according to the formula $D_N = 1-\Sigma \text{ (allelic frequency)}^2 * 100$ (Weir, 1990). Dendrograms were constructed with BioNumerics using a categorical multi-state coefficient and UPGMA (unweighted pair group method with arithmetic mean) clustering. Isolates having indistinguishable MLVA types were compared to PFGE data to determine the discriminatory capability of the developed non-O157 STEC MLVA method in comparison to PFGE.

RESULTS

Selection of VNTR loci

The first step in MLVA development is to search for and identify candidate VNTR loci in targeted genomes, which was accomplished in this study by using the free online Tandem Repeats Finder software (Benson, 1999). Several program parameters can be adjusted when searching for tandem repeats within a DNA sequence. Smith-Waterman style local alignment weights for matches, mismatches, and indels, the maximum period size to report, and the minimum alignment score to report can all be adjusted by the user. The alignment score is calculated according to the selected weights

for matches, mismatches, and indels among the tandem repeats comprising the VNTR array and the period size is the length of the entire VNTR array. Candidate VNTR loci for MLVA assays should have as few mismatches and indels within tandem repeat arrays as possible and have a high alignment scores, representing well-conserved VNTR loci.

In this study, the completed genomic sequences of 3 non-O157 STEC strains (O26, O103, and O111) were screened for candidate VNTR loci with Smith-Waterman alignment weights for matches, mismatches, and indels set at 2, 7, and 7, respectively. The minimum alignment score to report was set at 40, and maximum period size was set at 500. These parameters were chosen to narrow the number of reported tandem repeat arrays to those that had larger tandem repeat sequences (> 3 nt), larger copy numbers with minimal mismatching and indels, and period sizes less than 500 nt since capillary electrophoresis with a 600 nt size standard was to be used for sizing of PCR-amplified VNTR loci.

Interestingly, a comparison of short tandem repeat structures reported for 4 STEC O157:H7 strains, 2 generic *E. coli* strains, and 3 non-O157 STEC strains revealed that there are many more tandem repeats having a copy number greater than 5.0 in STEC O157:H7 strains than in all of the others, while the total number of tandem repeats reported were similar between STEC O157:H7 strains and non-O157 STEC strains (Table 5). When searching for candidate VNTR loci for strain discrimination, higher copy numbers within conserved tandem repeat arrays are desirable since tandem repeat arrays with lower copy numbers are likely to be less variable. As a result, fewer VNTR loci within non-O157 STEC strains were easily identifiable as candidates for MLVA as compared to those in STEC O157:H7 strains. The number of tandem repeats having

higher copy numbers among the non-O157 STEC strains was more similar to those found in 2 strains of generic *E. coli* K-12, which has an approximately 800 Kb smaller genome. A total of 12 VNTR loci, exhibiting differing levels of diversity among the genomic sequences of the 3 non-O157 STEC strains, were selected for further evaluation and characterization of their diversity among the big 6 non-O157 STEC serogroups (Table 4).

Characteristics of selected VNTR loci

Since the majority of bacterial genomes code for proteins, it was expected that most VNTR arrays would be located within genes. Of the 12 selected VNTR loci evaluated, 10 are located on the bacterial chromosome and 2 on plasmids. According to BLAST searches against the NCBI database, all chromosomal VNTR loci are located within sequences coding for known or putative proteins but the plasmid located VNTR loci had no known functions (Table 4). Of special interest was SVL-6, which had flanking sequences very similar to a *stx*2 converting phage, *stx*2 being one of the major virulence factors of STEC. In addition to being located within genes, VNTR loci would also be expected to occur in multiples of 3 so that the reading frame is not disrupted. It has been previously noted that a preponderance of 6 nt repeats exist throughout STEC O157:H7 genomes (Keys *et al.,* 2005). Additionally, it was hypothesized that tandem repeats that do not disrupt the open reading frame and are multiples of 3 are more likely to be of random origin than those that are not, resulting in more diversity. Accordingly, most of the selected loci were 6 nt in length or multiples of 3 (Table 4). Three of the 12 selected loci contained tandem repeats that were not multiples of 3 to see if less intra-serogroup diversity was observed in those cases.

Evaluation of selected VNTR loci and optimization of multiplex PCR

Initial screening of the 12 selected VNTR loci and VNTR-specific primer sets was done with 24 isolates from the non-O157 STEC reference set (Table 1). All primer sets, except SVL-11, allowed amplification of the target VNTR loci in at least 4 of the 6 non-O157 STEC serogroups. All loci were polymorphic, ranging from 2 to 11 alleles per locus (Table 6) and no 2 isolates of different serogroups had the same MLVA type. Compared to previous MLVA studies (Hyytia-Trees *et al.,* 2006; Lindstedt *et al.,* 2004; Boxrud *et al.,* 2007), a relatively high number of null alleles, a result of VNTR locus and flanking sequence absence or polymorphism at priming sites, were observed in serogroups O145, O45, and O121. This result was logical since only one strain each from serogroups O103, O111, and O26 were available for multiple sequence alignment during primer design. Nucleotide variations in priming sites may not have allowed optimal PCR amplification of serogroups O145, O45, and O121 at some loci. Yet, null alleles are still treated as alleles and can aid discrimination, particularly for inter-serogroup discrimination. Based on the ranges of band sizes in agarose gels for each of the 12 loci, 3 multiplex PCR reactions with fluorescently labeled forward PCR primers were tested and optimized for fragment analysis. Reaction 1 contained primer sets SVL-1, SVL-3, SVL-4, and SVL-8, reaction 2 contained primer sets SVL-2, SVL-6, SVL-10, and SVL-12, and reaction 3 contained primer sets SVL-5, SVL-7, SVL-9, and SVL-11. Optimization of peak heights in fragment analysis was accomplished by adjusting the concentrations of PCR primers in multiplex reactions.

The usefulness of the selected VNTR loci for MLVA typing of multiple serogroups of non-O157 STEC was further tested using 10 isolates from each of the 'big 6' serogroups that were obtained from CDC and consisted of epidemiologically related outbreak cases as well as non-epidemiologically related sporadic cases (Table 2). It was expected that less discrimination would be possible among the closely related outbreak associated isolates than among the sporadic isolates. A similar level of diversity was observed for each of the 12 loci when comparing both sets of isolates (Table 6). A moderate to high diversity index was observed for most of the selected loci. Among all 84 non-O157 STEC isolates, SVL-3 had the highest diversity index, followed by SVL-6 and SVL-8. SVL-1 was the most polymorphic locus with 18 different alleles (Table 7). However, SVL-1 had only a moderate diversity index (59.8) due to the lack of amplification of the locus in serogroups O121, O45, and O145. Loci SVL-6, SVL-3, and SVL-2 also exhibited high levels of polymorphism with 15, 13, and 9 alleles, respectively (Table 7). Only loci SVL-7, SVL-9, and SVL-11 had relatively low diversity indices at 47.6, 17.4, and 32.0, respectively (Table 7). Locus SVL-7 had the largest allelic range (1-38 tandem repeats) within serogroup O145 but had a low overall diversity. Locus SVL-9 had the lowest overall level of diversity and aided in discrimination only within serogroup O121. Locus SVL-11 was highly polymorphic only within serogroup O111, as was expected since this locus is plasmid located.

Several loci exhibited little or no intra-serogroup diversity but had distinct inter-serogroup diversity, helping discriminate between serogroups (Table 7). For example, locus SVL-4 contained 12 tandem repeats in all 14 O26 isolates, 9 tandem repeats in 13 of 14 O111 isolates, and 10 tandem repeats in 13 of 14 O121 isolates. Among the

individual serogroups, the characteristics of the 12 loci varied considerably and it may be possible to tailor the MLVA method for each serogroup by retaining the most diverse loci and discarding the least diverse. However, when typing all 6 serogroups simultaneously, discarding any of the 12 loci, except SVL-9, decreased the inter-serogroup discriminatory capability.

O26. The highest level of discriminatory capability was achieved in serogroup O26. All 14 isolates tested exhibited a unique MLVA type that differed from all other O26 isolates by 1 or more tandem repeats at 2 or more loci. Twelve different alleles were observed in locus SVL-1 alone. This high level of discrimination was achieved with just 4 loci (Table 7). Omitting all loci except SVL-1, SVL-2, SVL-3, and SVL-6 had no effect on the discriminatory capability. Therefore, a STEC O26-specific MLVA assay may be possible when loci SVL-1, SVL-2, SVL-3, and SVL-6 are targeted. Additionally, these 4 loci can be combined in a single multiplex PCR reaction using the existing fluorescent labels with no fragment overlap with the same label.

O111. Serogroup O111 had a low percentage of null alleles and the highest diversity indices, even though little or no diversity was observed in 7 loci (SVL-4, SVL-5, SVL-7, SVL-8, SVL-9, SVL-10, and SVL-12) and omitting these 7 loci had little effect on discrimination. The remaining 5 loci had a relatively high level of diversity, ranging from 74.7 to 90.1 (Table 7). A total of 11 unique O111 MLVA types were observed, with two groups of identical MLVA types. Isolates K6807, K6808, K6809, and K7091 were all from a single cluster of illnesses isolated in Oklahoma. Isolates K6807, K6809, and K7091 were indistinguishable by MLVA, but isolate K6808 had null alleles at loci SVL-11 and SVL-12. It is expected that among closely related isolates

MLVA types would differ by only a few tandem repeats at 1 or 2 loci, not by the absence of the entire VNTR region. It is more likely that a polymorphism within one or both of the priming sites at each locus accounted for the null alleles in isolate K6808. The second group of identical O111 MLVA types occurred between 2 isolates from an outbreak occurring in a Colorado prison.

O103. Serogroup O103 exhibited only moderate diversity. Only locus SVL-3 had a high diversity index at 87.5. Only loci SVL-3, SVL-4, SVL-5, and SVL-12 were required to provide the observed level of discrimination (Table 7). Three pairs of identical MLVA types were observed among 11 unique MLVA types for the 14 O103 isolates tested. The first pair of identical isolates, PT91-24 and TB154A, were from the non-O157 reference set from the STEC Center and both were isolated from Washington. However, the two strains were isolated a year apart and were of different serotypes (different H antigens). As a result, it is unlikely that an epidemiological connection existed. Isolates K3530-1 and K3529-1, isolated in association with goats in Nebraska, were indistinguishable by MLVA. The third pair of O103 isolates identical by MLVA, 2010C-3251 and 2010C-3219, both were isolated from Iowa and an epidemiological connection cannot necessarily be ruled out from the data provided by CDC.

O121. Only 3 loci exhibited moderate to high diversity in serogroup O121. Only SVL-3, SVL-6, and SVL-9 were needed to provide the observed level of discrimination (Table 7). Twelve unique MLVA types were observed among the 14 O121 isolates. One group of 3 isolates (K5313, K5316, and K5323) were indistinguishable by MLVA. All 3 strains were isolated from patients in Colorado and an epidemiological connection cannot be ruled out by the data provided by CDC.

O45. Low levels of diversity were observed for the 12 selected loci with serogroup O45. No PCR amplification of loci SVL-1, SVL-2, SVL-6, and SVL-11 occurred and although amplification was possible for loci SVL-7, SVL-8, SVL-9, SVL-10 and SVL-12, no allelic variation was observed (Table 7). Among the 14 isolates tested, 9 unique MLVA types were observed with 2 groups of identical MLVA types. The first group consisted of isolates K3472, 3095-04, and MI05-14, isolated from North Carolina, Maine, and Michigan, respectively. The second group consisted of isolates 3506-04, 3093-04, MI01-88, and DA-21, isolated from Michigan, Maine, Michigan, and Florida, respectively. No epidemiological connection is likely for each of these two groups.

O145. Low diversity indices were observed also for all 12 loci with serogroup O145. PCR amplification was not possible at locus SVL-11, and little or no diversity was observed for loci SVL-4, SVL-5, SVL-8, SVL-9, SVL-10, and SVL-12 (Table 7). Of 14 isolates tested, 8 MLVA types were observed with 2 groups of identical MLVA types. Isolates 4865/96 and GS-G5578620, isolated in Germany and Nebraska, respectively, were of different serotypes and had no logical epidemiological connection. The second group of identical O145 MLVA types consisted of isolates 2010C-3513, 2010C-3515, 2010C-3507, 2010C-3508, and 2010C-3526c1 and were all isolated from Michigan and Ohio and associated with a single cluster of illnesses.

Comparison with generic *E. coli*, EPEC, and STEC O157

Five STEC O157:H7 strains (C7927, EO144, F4546, K3995, and SEA-13B88), 2 EPEC strains (O119 and O55), and 1 strain of generic *E. coli* K-12 were MLVA typed

with the selected loci and compared to the MLVA types of the 24-isolate non-O157 STEC reference set (Table 3). All 8 strains exhibited unique MLVA types with low to moderate diversity indices, when amplification was possible (data not shown). Among the 5 O157:H7 isolates, no PCR amplification was possible for loci SVL-10 and SVL-12. Among the 2 EPEC isolates and K-12 strain, no PCR amplification was possible at loci SVL-1, SVL-2, SVL-6, SVL-10, SVL-5, and SVL-7. The same tandem repeat number was observed in all 8 isolates for locus SVL-9, located within a gene coding for a chaperonin protein. SVL-9 was also present in all 84 non-O157 STEC isolates and had the lowest diversity index of all loci. A minimum spanning tree of the 32 isolates used in this comparison grouped STEC O157:H7 isolates in a separate cluster (Figure 1).

Correlation of MLVA data with PFGE data for 36 CDC isolates

Twenty one unique MLVA types were observed among the 24-isolate non-O157 STEC reference set. Since each of the 4 isolates from the 6 serogroups were not epidemiologically related and original isolations were separated temporally, spatially, and sometimes by serotype, it was expected that each isolate would produce a unique MLVA type. However, 3 pairs of isolates were indistinguishable by MLVA—isolates MI01-88 and DA-2 (O45), TB154A and PT91-24 (O103), and 4865/96 and GS-G5578620 (O145). Interestingly, each pair of isolates was composed of different serotypes. Unfortunately, PFGE data is not available for the 24-isolate non-O157 STEC reference set from the STEC Center and comparison of MLVA data with PFGE data for these isolates was not possible.

A total of 65 unique MLVA types were identified among the 84 non-O157 STEC isolates tested. For the most part, serogroups clustered together in a minimum spanning tree (Figure 2) and by UPGMA (Figure 3). All serogroups differed from each other by 1 or more tandem repeats at 3 or more loci, except for among O103, O45, and one isolate of O111 (Figure 2). A similar or slightly better level of discrimination was observed when typing the isolates in this study by MLVA and PFGE, according to the information provided by CDC. For serogroup O26, the developed MLVA method was able to discriminate among two isolates (2009EL1049 and K5537) that had an identical *Xba*I PFGE pattern. MLVA types for serogroups O121 and O103 were in accordance with the PFGE data provided—two O103 isolates indistinguishable by MLVA were also indistinguishable by PFGE with *Xba*I and *Bln*I. In serogroup O111, 2 outbreak associated isolates (2010EL-1239 and 2010EL-1240) were distinguishable by PFGE with *Xba*I but not with *Bln*I or MLVA. Another group of outbreak associated isolates in serogroup O111, composed of 4 isolates (K6807, K6808, K6809, and K7091), were indistinguishable by PFGE with *Bln*I, but one isolate (K7091) had a different Xbal pattern. However, MLVA types for isolates K7091, K6807, and K6808 were indistinguishable and K6809 had a different MLVA type. In serogroups O145 and O45, a lower level of discrimination was possible than with PFGE. Two identical O45 isolates by MLVA had different PFGE patterns and 6 identical O145 isolates by MLVA had 4 different *Xba*I patterns and 3 different *Bln*I patterns.

DISCUSSION

In this study, a MLVA assay was developed for strain discrimination among the 6 most commonly isolated serogroups of non-O157 STEC. The assay was designed based on 12 VNTR loci and validated by fingerprinting a collection of 84 strains of the big 6 serogroups of non-O157 STEC. A total of 65 unique MLVA types were observed among 84 non-O157 STEC isolates (14 from each serogroup). The MLVA results clustered the epidemiologically related isolates together and differentiated the nonrelated isolates, also providing a discriminatory power similar to slightly better than that of PFGE. Discrimination by MLVA was greater in serogroup O26, similar in serogroups O103, O121, and O111, and decreased in serogroups O45 and O145 when compared to PFGE. The developed MLVA method was simple, rapid, and reproducible with easy-to-interpret and portable results. Closely mimicking the current CDC protocols for MLVA, the developed method should be easy to standardize, facilitating data exchange. Non-O157 STEC, as emerging pathogens, are being isolated more frequently in association with human illness (Johnson *et al.,* 2006; Coombes *et al.,* 2008), and a rapid and highly discriminatory subtyping technique, such as the MLVA method developed in this study, would be valuable to augment PFGE in outbreak investigations.

Since multiple serogroups were targeted in this study, potentially highly diverse VNTR loci were chosen to aid in intra-serogroup discrimination and potentially less diverse VNTR loci were chosen to aid in inter-serogroup discrimination. All chromosomally located VNTR loci were contained within DNA sequences coding for known or putative proteins (Table 4). Tandem repeat arrays that are located within genes and have repeat lengths in multiples of 3, therefore not altering the open reading frame,

are likely to be more conserved than those located outside of gene sequences (Keys *et al.,* 2005). We speculate that while the number of tandem repeats may be highly variable within genes, the tandem repeat array and flanking sequences within genes are likely to be present in most serogroups and this feature was an important consideration when selecting VNTR loci in this study. Three of the selected VNTR loci (SVL-7, SVL-8, and SVL-10) contained tandem repeats that were not multiples of 3. It was expected that a lower level of diversity would be observed with these loci. As expected, the 3 loci exhibited a low to moderate intra-serogroup diversity but a moderate to high inter-serogroup diversity. The rate of tandem repeat loss or addition among tandem repeat lengths that are not multiples of 3 is likely less frequent and helped discriminate among serogroups in this study. The locus having the lowest overall diversity index (SVL-9) had a tandem repeat length of 9 and was located within a gene coding for a chaperonin protein, suggesting that less tandem repeat variation may be favored within more conserved genes coding for proteins that have critical structure and function. Two of the selected VNTR loci were plasmid located to allow for evaluation of the diversity among multiple serogroups. Locus SVL-12 was amplified by PCR in all non-O157 STEC serogroups except O145 and both EPEC strains, but was not amplified in O157:H7 or *E. coli* K-12, and only a moderate level of diversity was observed. PCR amplification of locus SVL-11, on the other hand, was possible only in serogroup O111 and one isolate of O103. SVL-11 was highly diverse among O111 and was useful in strain discrimination among this serogroup. VNTR loci located on plasmids, if present, may serve as useful serogroup identifiers when multiple serogroups are targeted in a single MLVA assay.

The genomic locations of loci SVL-6 were of special interest. Based on a BLAST search against the NCBI database, SVL-6 was located within a gene sharing high homology to a *stx*2 converting phage (Smith *et al.,* 2012). Stx2 is one of the major virulence factors of STEC and is frequently associated with the development of HUS (Nataro and Kaper, 1998). It is believed that the *stx*2 gene can be acquired by STEC following contact with *stx*2 converting phages and subsequent incorporation of the sequence into previously non-pathogenic or less pathogenic *E. coli* genomes (Scheutz *et al.,* 2011). As expected, SVL-6-specific PCR primers were able to generate amplicons in serogroups O157, O26, O111, O103, O121, and O145—the serogroups most commonly associated with Shiga toxin production—but not in 2 EPEC strains or in *E. coli* K-12. A recent major outbreak of human illness was caused by an EAEC strain that acquired the *stx*2 gene, giving it an increased level of virulence (Scheutz *et al.,* 2011). Non-O157 STEC are difficult to distinguish from non-pathogenic *E. coli* by cultural methods and detection of non-O157 STEC in the clinical setting is often a challenge (Werber *et al.,* 2008). Candidate non-O157 STEC colonies are often selected for PCR-based detection of virulence genes prior to serological confirmation. By including a VNTR region within a gene having high homology to an *stx*2 converting phage in a non-O157 STEC MLVA assay, it may be possible to identify isolates with the potential to express Stx2, further validating an isolate as non-O157 STEC even before serological or biotyping data may be available. While MLVA methods specifically targeting virulence genes have been developed for other bacterial pathogens, more VNTR loci located within virulence associated genes may be especially useful for non-O157 STEC. Additionally,

SVL-6 exhibited a high level of overall diversity among the 84 non-O157 STEC isolates tested.

A total of 65 unique MLVA types were observed among 84 non-O157 STEC isolates (14 isolates from each serogroup), all of which were clinical isolates associated with outbreaks or sporadic cases. All serogroups had distinctly different MLVA types that differed by at least 3 loci, except for a few isolates of O103 and O45, which differed by one locus (Figure 2). For the most part, serogroups clustered together in a minimum spanning tree and most identical MLVA types were composed of epidemiologically related isolates. Less serogroup clustering was observed in the dendrogram by UPGMA (Figure 3), which was expected since MLVA is not well suited for evolutionary or phylogenetic studies (Lindstedt *et al.,* 2007), especially when multiple serogroups are evaluated. Genome-wide subtyping methods that are based on genetic changes that occur more slowly, such as PFGE, MLST, and multi-locus enzyme electrophoresis (MLEE), are better suited for studies of evolutionary relatedness. The MLVA assay developed in this study was successful in that it allowed inter- and intra-serogroup discrimination of 6 serogroups of non-O157 STEC.

The developed MLVA assay was more useful in fingerprinting some serogroups than others. High discriminatory power was observed for multiple loci in serogroups O26 and O111, and moderate to high discriminatory power for serogroups O103 and O121. The Simpson's diversity indices for these serogroups were comparable to other MLVA assays that have been proven to be useful in outbreak investigations, providing higher strain discrimination than with PFGE alone (Hyytia-Trees *et al.,* 2006; Lindstedt *et al.,* 2004; Boxrud *et al.,* 2007). Among serogroups O26, O111, O103, and O121, only

one pair of identical MLVA types (PT91-24 and TB154A) had no logical epidemiological connection. Most of the discriminatory power for these serogroups was provided by only 4 to 5 loci and development of a MLVA assay with 1 multiplex PCR reaction for serogroups O26, O111, and O111 may be possible with VNTR loci SVL-1, SVL-2, SVL-3, and SVL-6. A lower level of diversity was observed with the selected loci among serogroups O145 and O45. Several groups of identical MLVA types were identified among serogroups O145 and O45 composed of isolates that had no epidemiological connection, suggesting that the VNTR loci selected in this study may not be well suited for epidemiological investigation of serogroups O145 and O45. These results were expected since the only genomic sequences available for evaluation and selection of candidate VNTR loci were strains of serogroups O26, O111, and O103. A high percentage of null alleles were observed when MLVA typing serogroups O121, O45, and O145, due to absence of the VNTR loci or sub-optimal PCR primer binding. The major limiting factor for developing a more discriminatory method is the lack of available non-O157 STEC genomes for analysis. *In silico* comparison of multiple closely related strains would allow selection of the most polymorphic VNTR loci for the development of optimal MLVA assays. More non-O157 STEC genomes may be available soon as a result of the 100 K Genome Project, a collaborative effort between government, academia, and industry to sequence 100,000 human pathogenic foodborne bacteria.

Comparison of the MLVA assay designed for non-O157 STEC with 5 STEC O157:H7 isolates, 2 EPEC isolates, and 1 isolate of generic *E. coli* K-12 revealed that the MLVA method developed in this study is not very useful for typing these groups of *E.*

coli. While 8 unique MLVA types were observed for the 8 isolates, PCR amplification was not possible for many of the loci and several non-specific fragments were observed for loci SVL-8, SVL-5, and SVL-7 when amplification was possible. The same tandem repeat number was observed in all 8 isolates for locus SVL-9, located within a gene coding for a chaperonin protein. Additionally, SVL-9 as present in all 84 non-O157 STEC isolates tested and had the lowest overall diversity index of all 12 loci (Table 7). This gene and protein may be well conserved throughout *E. coli* and may serve as a useful identifier if unknown isolates are typed by the developed MLVA assay. Only *E. coli* and *Shigella* had high alignment scores in a BLAST search against the NCBI database with the SVL-9 flanking sequences.

In conclusion, the MLVA method for the 6 major non-O157 STEC serogroups developed in this study provided an expected level of inter- and intra-serogroup strain discrimination, given that only genome sequences of O26, O111, and O103 were available during assay development. Several highly diverse loci were identified that aided in discrimination of clinical non-O157 STEC isolates. Subsets of the 12 selected loci may be individually tailored for each serogroup to provide a high level of intra-serogroup discrimination. However, a majority of the selected loci are needed for a highly discriminatory inter-serogroup MLVA assay. Until more genome sequences are available for comparison, a higher overall level of discrimination may not be possible. Further validation of the developed method with more isolates from outbreak and sporadic cases is needed and should be compared to PFGE data for all isolates to gain a more complete understanding of the usefulness of this method for intra- and inter-serogroup discrimination of epidemiologically related non-O157 STEC isolates.

LITERATURE CTITED

Besser, R.E., P.M. Griffin, and L. Slutsker. (1999). *Escherichia coli* O157:H7 gastroenteritis and the hemolytic uremic syndrome: an emerging infectious disease. *Annu Rev Med*. 50:355–67.

Bettelheim, K.A. (2007). The Non-O157 Shiga-Toxigenic (Verocytotoxigenic) Escherichia coli; Under-Rated Pathogens. *Critical Reviews in Microbiology/ 33*(1):67-87.

Boxrud, D., K. Pederson-Gulrud, J. Wotton, C. Medus, E. Lyszkowicz, J. Besser, and J. M. Bartkus. (2007). Comparison of Multiple-Locus Variable-Number Tandem Repeat Analysis, Pulsed-Field Gel Electrophoresis, and Phage Typing for Subtype Analysis of *Salmonella enterica* serotype Enteritidis. Journal of Clinical Microbiology. 45(2):536-543

Brooks, J.T., E.G. Sowers, J.G. Wells, K.D. Greene, P.M. Griffin, R.M. Hoekstra, and N.A. Strockbine. (2005). Non-O157 Shiga Toxin–Producing Escherichia coli Infections in the United States, 1983–2002. *Journal of Infectious Diseases.* 192(8):1422-1429.

Centers for Disease Control and Prevention (CDC). (2012). National Shiga toxin-producing *Escherichia coli* (STEC) Surveillance Overview. Atlanta, Georgia: US Department of Health and Human Services.

Centers for Disease Control and Prevention (CDC). (2012b). National Shiga toxin-producing *Escherichia coli* (STEC) Surveillance Annual Summary, 2009. Atlanta, Georgia: US Department of Health and Human Services.

Centers for Disease Control and Prevention (CDC). (2012c). National Shiga toxin-producing *Escherichia coli* (STEC) Surveillance Annual Summary, 2008. Atlanta, Georgia: US Department of Health and Human Services.

Cho, S., D.J. Boxrud, J.M. Bartkus, T.S. Whittam, and M. Saeed. (2007). Multiple-locus variable-number tandem repeat analysis of *Salmonella* Enteritidis isolates from human and non-human sources using a single multiplex PCR. FEMS Microbiology Letters. 275 (1):16–23.

Cho, S., F. Diez-Gonzalez, C.P. Fossler, S.J. Wells, C.W. Hedberg, J.B. Kaneene, P.L. Ruegg, L.D. Warnick, and J.B. Bender. (2006). Prevalence of shiga toxin-encoding bacteria and shiga toxin-producing Escherichia coli isolates from dairy farms and county fairs. *Vet Microbiol.* 118(3-4):289-98.

Coombes, B.K., M.E. Wickham, M. Mascarenhas, S. Gruenheid, B.B. Finlay, and M.A. Karmali. (2008). *Appl Environ Microbiol.* 74(7): 2153–2160.

Cray, W.C. Jr, L.A. Thomas, R.A. Schneider, and H.W. Moon. (1996). Virulence attributes of Escherichia coli isolated from dairy heifer feces. *Vet Microbiol.* 53(3-4):369-74.

Donnenberg, M.S. and T.S. Whittam. (2001). Pathogenesis and evolution of virulence in enteropathogenic and enterohemorrhagic Escherichia coli. *The Journal of Clinical Investigation.* 107(5): 539-548.

Ethelberg, S., K. Olsen, S. Flemming, C. Jensen, P. Schiellerup, J. Engberg, A. Munk Petersen, B. Olesen, P. Gerner-Smidt, and K. Molbak. (2004). Virulence factors for hemolytic uremic syndrome, Denmark. *Emerg Infect Dis*. Available at: http://wwwnc.cdc.gov/eid/article/10/5/03-0576.htm.

Gyles, C.L. (2007). Shiga toxin-producing *Escherichia coli*: an overview. *J Anim Sci.* 85(Suppl 1):E45–E62.

Hundy, R.L. and S. Cameron. (2004). Risk factors for sporadic human infection with shiga toxin-producing *Escherichia coli* in South Australia. *Commun Dis Intelligence* 28:74–79.

Johnson, K.E., C.M. Thorpe, and C.L. Sears. (2006). The Emerging Clinical Importance of Non-O157 Shiga Toxin—Producing Escherichia coli. *Clinical Infectious Diseases.* 43(12):1587-1595.

Hyytiä-Trees, E., S.C. Smole, P.A. Fields, B. Swaminathan, and E.M. Ribot. (2006). Second generation subtyping: A proposed PulseNet protocol for multiple-locus variable-number tandem repeat analysis of Shiga toxin–producing *Escherichia coli* O157 (STEC O157). *Foodborne Pathog. Dis.* 3:118-131

Johnson, R.P., R.C. Clarke, and J.B. Wilson JB. (1996). Growing concerns and recent outbreaks involving non-O157:H7 serotypes of verotoxigenic *Escherichia coli*. *J Food Prot*. 59:1112–22.

Kaplinski, L., R. Andreson, T. Puurand, and M. Remm. (2005). MultiPLX: automatic grouping and evaluation of PCR primers. *Bioinformatics*. 21(8):1701-1702.

Karama, M. and C.L. Gyles. (2010). Methods for Genotyping Verotoxin-Producing Escherichia coli. *Zoonoses and Public Health.* 57(7-8): 447-462.

Karmali, M.A., M. Mascarenhas, S. Shen, K. Ziebell, S. Johnson, R. Reid-Smith, and J.B. Kaper. (2003). Association of Genomic O Island 122 of Escherichia coli EDL 933 with Verocytotoxin-Producing Escherichia coli Seropathotypes That Are Linked to Epidemic and/or Serious Disease. *Journal of Clinical Microbiology.* 41(11): 4930-4940.

Kaspar, C., M. E. Doyle, and J.Archer. (2009). White Paper on Non-O157:H7 Shiga Toxin-Producing *E. coli* from Meat and Non-Meat Sources. Food Research Institute, UW–Madison, Dec. 2009/April 2010

Keim, P., L.B. Price, A.M. Klevytska, K.L. Smith, J.M. Schupp, R. Okinaka, and M.E. Hugh-Jones. (2000). Multiple-Locus Variable-Number Tandem Repeat Analysis Reveals Genetic Relationships within Bacillus anthracis. *Journal of Bacteriology.* 182(10): 2928-2936.

Lindstedt, B.A., E. Heir, E. Gjernes, and G. Kapperud. (2003). DNA fingerprinting of *Salmonella enterica* subsp. *enterica* serovar Typhimurium with emphasis on phage type DT104 based on the variable number tandem repeat loci. *J. Clin. Microbiol.* 41:1469-1479.

Lindstedt, B.A., T. Vardund, L. Aas, and G. Kapperud. (2004). Multiple-locus variable-number tandem-repeats analysis of *Salmonella enterica* subsp. *enterica* serovar Typhimurium using PCR multiplexing and multicolor capillary electrophoresis. *J. Microbiol. Method.* 59:163-172.

McPherson, M., M. Kirk, J. Raupach, and A. Koehler. (2008). Risk factors of sporadic human infection of shiga toxin producing *Escherichia coli* O157 and non-O157 in Australia. *Int J Infect Dis.* 12:E217.

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

Nielsen, E.M., F. Scheutz, and M. Torpdahl. (2006). Continuous surveillance of Shiga toxin-producing *Escherichia coli* infections by pulsed-field gel electrophoresis shows that most infections are sporadic. *Foodborne Path Dis*. 3:81–87.

O'Brien, A.D., V.L. Tesh, A. Donohue-Rolfe, M.P. Jackson, S. Olsnes, K. Sandvig, A.A. Lindberg, and G.T. Keusch. (1992). Shiga toxin: biochemistry, genetics, mode of action, and role in pathogenesis. *Curr Top Microbiol Immunol.* 180:65--94.

Ogura, Y., T. Ooka, A. Iguchi, H. Toh, M. Asadulghani, K. Oshima, T. Kodama, H. Abe, K. Nakayama, K. Kurokawa, T. Tobe, M. Hattori, and T. Hayashi. (2009). Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli. Proc Nat Acad Sci* USA 106:17939–17944.

Renter, D.G., J.G. Morris Jr, J.M. Sargeant, L.L. Hungerford, J. Berezowski, T. Ngo, K. Williams, and D.W. Acheson. (2005). Prevalence, risk factors, O serogroups, and virulence profiles of Shiga toxin-producing bacteria from cattle production environments. *J Food Prot.* 68(8):1556-65.

Rhoades, J.R., G. Duffy, and K. Koutsoumanis. (2009). Prevalence and concentration of verocytotoxigenic *Escherichia coli, Salmonella Enterica* and *Listeria monocytogenes* in the beef production chain: a review. *Food Microbiol*. 26:357–376.

Ribot, E.M., M.A. Fair, R. Gautom, D.N. Cameron, S.B. Hunter, B. Swaminathan, and T.J. Barret. (2006). Standardization of pulsed-field gel electrophoresis protocols for the subtyping of *Escherichia coli* O157:H7, *Salmonella*, and *Shigella* for PulseNet. *Foodborne Pathogens and Disease*. 3:59–67.

Rivas, M., S. Sosa-Estani, J. Rangel, M.G. Caletti, P. Vallés, C.D. Roldán, L. Balbi, M.C.M. De Mollar, D. Amoedo, E. Miliwebsky, I. Chinen, R.M. Hoekstra, P. Mead, and P.M. Griffin. (2008). Risk factors for sporadic Shiga toxin-producing *Escherichia coli* infections in children. Argentina. *Emerg Infect Dis* 14:763–771.

Schimmer, B., K. Nygaard, H. Eriksen, J. Lassen, B.A. Lindstedt, L.T. Brandal, G. Kapperud, and P. Aavitsland. (2008). Outbreak of haemolytic syndrome in Norway caused by stx_2-postitive *Escherichia coli* O103:H25 traced to cured mutton sausage. *BMC Inf Dis*. 8.

"Shiga Toxin-Producing Escherichia coli in Certain Raw Beef Products." Federal Register. 77: 105 (31 May 2012). Department of Agriculture. Food Safety and Inspection Service.

Smith, D., D. Rooks, P. Fogg, A. Darby, N. Thomson, A. McCarthy, and H. Allison. (2012). Comparative genomics of Shiga toxin encoding bacteriophages. *BMC Genomics.* 13(1): 311.

Swaminathan, B., T.J. Barrett, S.B. Hunter, and R.V. Tauxe. (2001). PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerg. Infect. Dis.* 7:382–389.

Tarr, P.I., C.A. Gordon, and W.L. Chandler. (2005). Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. *Lancet.* 365:1073–1086

Todar, K. (2011). "Pathogenic *E. coli*". Online Textbook of Bacteriology. University of Wisconsin–Madison Department of Bacteriology. Available at: http://www.textbookofbacteriology.net/e.coli.html. Accessed 19 May 2011.

Wahl, E., L. Vold, B. Lindstedt, T. Bruheim, and J. Afset. (2011). Investigation of an Escherichia coli O145 outbreak in a child day-care centre - extensive sampling and characterization of eae- and stx1-positive E. coli yields epidemiological and socioeconomic insight. *BMC Infectious Diseases.* 11(1):238.

Weir, B.S. (1990*).* Genetic Data Analysis. Sunderland, MA: Sinauer Associates.

Wells, J.G., L.D. Shipman, K.D. Greene, E.G. Sowers, J.H. Green, D.N. Cameron, F.P. Downes, M.L. Martin, P.M. Griffin, and S.M. Ostroff. (1991). Isolation of Escherichia coli serotype O157:H7 and other Shiga-like-toxin-producing E. coli from dairy cattle. *J Clin Microbiol.* 29(5):985-9.

Werber, D., C. Frank, M. Wadl, H. Karch, A. Fruth, and K. Stark. (2008). Looking for tips to find icebergs—surveillance of haemolytic uraemic syndrome to detect outbreaks of Shiga toxin-producing *E. coli* infection. *Euro Surveill* 13(9):pii=8053.

TABLES

Table 1. The 24 strains in the non-O157 STEC reference set.*

O	H	Host	Isolation Location	Isolation Date	Clinical Manifestation
26	11	Human	Australia (Brisbane)	1986	Diarrhea (bloody, acute)
26	11	Human (F, 2y)	USA (Idaho)	1997	HUS (expired)
26		Human	USA (Mont.)	1999-2000	
26	N	Human (child, 6y)	USA (Wash.)	1991	Diarrhea (chronic)
45	2	Human (M, 45y)	USA (Mich.)	2003	
45	2	Human (F, 38y)	USA (Mich.)	2001	
45	2	Human (M, 12y)	USA (Mich.)	2006	
45	NM	Human (F, 77y)	USA (Fla.)	1999	Diarrhea (bloody)
103	2	Human	USA (Mont.)	1999-2000	
103	6	Human	USA (Wash.)	1991	Diarrhea
103	25	Human (F, 3y)	USA (Idaho)		
103	N	Human	USA (Wash.)	1990	
111	2	Human (child)	France	1992	HUS (outbreak)
111	8	Human (F, 18y)	USA (TX)	1999	HC (outbreak)
111	11	Human	USA (Conn.)	2003	
111	NM	Human (M, 67y)	USA (Neb.)	1985	
121	19	Human (F, 51y)	USA (Mich.)	2000	
121	19	Human	USA (Mont.)	1998	
121		Human	USA (Mont.)	1999-2000	
121	[19]	Human (F, 51y)	USA (Mass.)	1998	Diarrhea (bloody)
145	16	Human	Canada	1987	HC (HUS)
145	[28]	Human	Germany	1996	HUS
145	NM	Human	USA (Neb.)	1998	Diarrhea
145	NT	Human	Uruguay		

*Information provided by the STEC Center of Michigan State University

Table 2. 60 non-O157 STEC isolates from CDC.

Serogroup	Isolate	Serotype	Isolation Location	Epidemiological Information	XbaI Pattern	BlnI pattern
O145	2010C-3517	O145:NM	MI	cluster 1004MIENM-1	ENMX01.0025	ENMA26.0018
	2010C-3515	O145:NM	MI	cluster 1004MIENM-1	ENMX01.0016	ENMA26.0017
	2010C-3507	O145:NM	OH	cluster 1004MIENM-1	ENMX01.0016	ENMA26.0017
	2010C-3508	O145:NM	OH	cluster 1004MIENM-1	ENMX01.0016	ENMA26.0017
	2010C-3526c1	O145:NM	MI	cluster 1004MIENM-1	ENMX01.0043	ENMA26.0018
	K6208	O145:NM	ND	STEC Broth		
	2011EL-1210	O145:NM	FL	STEC Broth		
	3060-04	O145:NM	UT	STEC Broth	ENMX01.0003	
	2010C-3513	O145:NM	MI	STEC Broth	ENMX01.0113	ENMA26.0078
	K2387	O145:NM	MD	STEC Broth	ENMX01.0030	
O111	K6807	O111:NM	OK	0808OKEXD-1	EXDX01.0050	EXDA26.0174
	K6808	O111:NM	OK	0808OKEXD-1	EXDX01.0050	EXDA26.0174
	K6809	O111:NM	OK	0808OKEXD-1	EXDX01.0050	EXDA26.0174
	K7091 CHROM 1	O111:H8	OK	0808OKEXD-1	EXDX01.0188	EXDA26.0174
	K5652 SMAC 11	O111:NM	IN	STEC Broth		
	2009EL1340	O111:NM	FL	STEC Broth		
	2010EL-1239	O111:NM	CO	1005COEXD-1	EXDX01.0123	EXDA26.0077
	2010EL-1240	O111:NM	CO	1005COEXD-1	EXDX01.0130	EXDA26.0077
	2010EL-2219	O111:H8	FL	STEC Broth		
	2010EL-2231	O111:H8	FL	STEC Broth		
O26	2009EL1049	O26:H11	OK	STEC Broth	EVCX01.0018	
	2011EL-1012	O26:H9	IN	STEC Broth		
	2011EL-1138	O26:H11	AK	STEC Broth		
	2010EL-1372	O26:NM	WA	Daycare OB		
	2009EL-1480	O26:H11	FL	STEC Broth	EVCX01.0676	
	2011EL-1233	O26:H11	NV	STEC Broth		
	2010EL-2220	O26:H11	FL	STEC Broth		
	K3621	O26:H11	CO	STEC Broth		
	K3651	O26:H11	NC	STEC Broth		
	K5537	O26:H11	MO	STEC Broth	EVCX01.0018	EVCA26.0218
O45	05-3031	O45:H2	UT	STEC Broth	EH2X01.0133	
	03-3300	O45:H2	MO	STEC Broth	EH2X01.0002	
	K3472	O45:H2	NC	STEC Broth		
	K3523	O45:H2	FL	STEC Broth		
	3506-04	O45:H2	MI	STEC Broth	EH2X01.0005	
	3001-04	O45:H2	MO	STEC Broth	EH2X01.0008	
	3065-04	O45:H2	WI	STEC Broth	EH2X01.0013	
	3093-04	O45:H2	MA	STEC Broth	EH2X01.0014	

	3095-04	O45:H2	MA	STEC Broth	EH2X01.0016	
	3105-04	O45:H2	MI	STEC Broth	EH2X01.0017	
O103	2009EL1342	O103:NM	FL	STEC Broth	EXWX01.0537	
	2009EL1295	O103:H2	IN	STEC Broth	EXWX01.0540	
	3546-05	O103:H25	VA	STEC Broth	EXWX01.0146	
	3409-05	O103:H25	VA	STEC Broth	EXWX01.0145	
	K3530-1	O103:H2	NE	Goat		
	K3529-1	O103:H2	NE	Goat		
	K3435	O103:H2	MO	STEC Broth		
	2010C-3251	O103:H2	IA	STEC Broth		
	2010C-3219	O103:H2	IA	STEC Broth		
	2009EL-1899	O103:H2	FL	STEC Broth		
O121	K5363	O121:H19	CT	STEC Broth		
	K5316	O121:H19	CO	STEC Broth		
	K5313	O121:H19	CO	0707COEXK-1	EXKX01.0001	EXKA26.0001
	K5223	O121:H19	CO	0707COEXK-1	EXKX01.0001	EXKA26.0001
	K3673	O121:H19	FL	STEC Broth		
	K3663	O121:H19	CO	STEC Broth	EXKX01.0013	EXKA26.0002
	K2126	O121:H19	VT	STEC Broth	EXKX01.0041	
	3294-06	O121:H19	WY	STEC Broth	EXKX01.0005	EXKA26.0001
	3326-06	O121:H19	NY	STEC Broth	EXKX01.0048	EXKA26.0002
	K2225	O121:H19	FL	STEC Broth	EXKX01.0044	

Table 3. *E. coli* isolates used for comparison.

E. coli Group	Strain	Epidemiological Information
STEC O157:H7	K3995	Spinach outbreak isolate
	C7927	Apple cider outbreak isolate
	F4546	Sprout-related outbreak isolate
	EO144	Meat isolate
	SEA-13B88	Clinical isolate from unpasteurized apple juice
EPEC	O119:H6	
	O55:H6	
Generic	K-12	

Table 4. Characteristics of the 12 VNTR loci used in this study.

Locus name	Repeat length (bp)	Location in O111:H- (5' end)	Primers (5'-3')	Primer T_m (°C)	Offset size in O111:H- (bp)	Primer Concentraion (µM)	Function
SVL-1	6	250070	F: **6FAM**-ACCCAGCCATTTCGTCAG R: ACGCAGATACCGTGGAG	59.63 61.65	251	0.05	Putative ATP-dependent Clp proteinase ATP-binding chain
SVL-2	6	2913106	F: **PET**-ACGCAGATACCGTGGAG R: TCAGGAATGTGGTGGTCTGTT	60.87 60.42	244	0.05	Hypothetical protein
SVL-3	6	4662685	F: **VIC**-TGGCAAACAGCACTACCATC R: GGACCAGTTAAGCCAGCAAA	59.72 60.25	248	0.04	Predicted protoheme IX synthesis protein HemY, Predicted uroporphyrinogen III methylase
SVL-4	15	810131[a]	F: **PET**-GGAAGAAGCAGCGAAGAAAG R: CATCGGGTGCCAGTTTTATG	59.34 61.27	270[a]	0.06	Membrane anchored protein TolA in TolA-TolQ-TolR complex
SVL-5	6	3051096[a]	F: **VIC**-GTCGTCTGTGGGATGCTCAA R: CAGCAATAACAGCAGGACGA	62.27 60.01	156[a]	0.05	Hydrogenase 4, Membrane subunit HyfF
SVL-6	9	2922513[a]	F: **6FAM**-GCAAGGGAAGTGGACAAA R: CTCCCATCGTTTCTGTTTCC	57.10 59.53	110[a]	0.07	Putative adenine methylase, Putative integrase, *stx*2 converting phage
SVL-7	8	774775	F: **6FAM**-CCTGGAAGCAGACAAATAACC R: GGACAATGCTACCGCCATAC	58.71 60.36	88	0.05	RhsC core protein with extension
SVL-8	13	1865598	F: **NED**-TACAGAGCGCGAGAAACAGA R: CCTGCCACCATCTCTTCAC	59.89 57.89	102	0.06	RhsE, IS677, Putative H repeat-associated protein
SVL-9	9	5160787	F: **PET**-CGTCGTCCGTGTCTGAATCT R: CAACCGAAGAATACGGCAAC	61.28 60.50	270	0.05	Chaperonin Cpn60
SVL-10	7	3346927	F: **VIC**-TTTGATGCAATGGTGGAGTG R: CACAAAGTGAGAGTCCGAAAA	60.52 57.99	166	0.05	Putative integrase
SVL-11	6[b]	35162[b]	F: **NED**-ATTCTGCTGTGGGCTTCTGT R: AATCAGAGCGGCAGGAAAA	59.87 60.87	90[b]	0.05	Plasmid located, no known function
SVL-12	9[c]	52289[c]	F: **NED**-CCGCAAGGGAAGCAGAAG R: TGCTGTTCCATCTCTTCTTCC	62.02 59.42	197[c]	0.04	Plasmid located, no known function

[a]Location or estimated amplicon size in *E. coli* O103:H2 strain 12009 (NC_013353.1)
[b]Location or estimated amplicon size in *E. coli* O111:H- str. 11128 plasmid pO111_3 (NC_013366.1)
[c]Location or estimated amplicon size in *E. coli* O26:H11 strain 11368 plasmid po26_2 (NC_013362.1)

Table 5. Comparison of genome size, number of reported tandem repeat arrays, and number of tandem repeat arrays with copy numbers greater than 5.0, according to Tandem Repeats Finder software, for 4 STEC O157:H7, 3 non-O157 STEC, and 2 *E. coli* K-12 strains.

***E. coli* strain**	**GenBank accession number**	**Genome size (bp)**	**Total number of tandem repeats[a]**	**Number of tandem repeats with copy number ≥ 5.0**
O157:H7 EC4115	NC_011353.1	5572075	302	20
O157:H7 TW14359	NC_013008.1	5528136	296	25
O157:H7 EDL933	NC_002655.2	5528445	285	23
O157:H7 Sakai	NC_002695.1	5498450	285	21
O111:H- 11128	NC_013364.1	5371077	242	9
O26:H11 11368	NC_013361.1	5697240	261	12
O103:H2 12009	NC_013353.1	5449314	339	9
K-12 DH10B	NC_010473.1	4686137	175	7
K-12 W3110	NC_007779.1	4646332	167	9

[a] As reported by Tandem Repeats Finder software with custom parameter settings.

Table 6. Comparison of VNTR loci characteristics between 60 clinical non-O157 STEC isolates and 24 isolates of a non-O157 STEC reference set.

	60 non-O157 STEC from CDC											
	R1				**R2**				**R3**			
	SVL-1	**SVL-3**	**SVL-4**	**SVL-8**	**SVL-2**	**SVL-6**	**SVL-10**	**SVL-12**	**SVL-5**	**SVL-7**	**SVL-9**	**SVL-11**
Fragment range (nt)	303-444	276-365	387-452	155-209	259-295	124-305	175-282	210-289	169-210	98-392	306-315	105-160
No. of alleles	14	11	5	5	7	10	5	4	7	4	2	8
Null alleles (%)	63.3	0	6.6	18.3	51.6	45.0	33.3	55.0	1.6	1.6	0	80.0
Allelic range	9-32	5-20	8-12	4-8	3-9	2-22	1-16	2-10	2-9	1-38	4-5	3-14
Diversity index	59.4	89.5	63.6	69.4	67.5	77.7	53.6	53.2	53.4	51.4	15.5	35.6
	24 non-O157 STEC from MSU											
	R1				**R2**				**R3**			
	SVL-1	**SVL-3**	**SVL-4**	**SVL-8**	**SVL-2**	**SVL-6**	**SVL-10**	**SVL-12**	**SVL-5**	**SVL-7**	**SVL-9**	**SVL-11**
Fragment range (nt)	309-400	270-371	418-452	155-259	253-283	133-305	175-181	211-307	169-186	98-106	306-315	147-153
No. of alleles	10	10	3	4	7	11	3	5	5	3	2	3
Null alleles (%)	62.5	0	0	29.2	58.3	54.2	37.5	45.8	4.2	8.3	0	87.5
Allelic range	10-25	4-21	9-12	4-12	2-9	3-23	1-2	2-12	2-5	1-2	4-5	10-11
Diversity index	62.0	86.6	47.5	69.9	64.9	71.4	54.0	63.8	54.3	36.6	22.8	23.6

Table 7. Comparison of VNTR loci characteristics for 6 non-O157 STEC serogroups.

O26

	R1				R2				R3			
	SVL-1	**SVL-3**	**SVL-4**	**SVL-8**	**SVL-2**	**SVL-6**	**SVL-10**	**SVL-12**	**SVL-5**	**SVL-7**	**SVL-9**	**SVL-11**
Fragment range	322-444	288-318	451-452	155-156	259-295	124-160	175-176	289	169 -170	97-98	306-307	132
Number of alleles	12	6	1	1	6	8	1	2	1	1	1	2
Null alleles (%)	0	0	0	0	7.1	0	0	92.9	0	0	0	92.9
Allelic range (TR)	12--31	7--12	12	4	3-9	2-6	1	10	2	1	4	7
Diversity index	97.8	86.8	0	0	83.5	82.4	0	14.3	0	0	0	14.3

O111

	R1				R2				R3			
	SVL-1	**SVL-3**	**SVL-4**	**SVL-8**	**SVL-2**	**SVL-6**	**SVL-10**	**SVL-12**	**SVL-5**	**SVL-7**	**SVL-9**	**SVL-11**
Fragment range	322-365	299-365	407-408	167-168	265-283	124-296	175-182	210-211	169-170	106-107	306-307	135-160
Number of alleles	7	8	2	2	6	4	3	3	1	2	2	6
Null alleles (%)	7.1	0	0	7.1	14.3	35.7	7.1	7.1	0	0	0	14.3
Allelic range	12-25	9-20	9	5	3-6	2-20	1-2	2	2	1-2	4-5	8-12
Diversity index	84.6	90.1	14.3	14.3	79.1	74.7	47.3	27.5	0	14.3	26.4	74.7

O103

	R1				R2				R3			
	SVL-1	**SVL-3**	**SVL-4**	**SVL-8**	**SVL-2**	**SVL-6**	**SVL-10**	**SVL-12**	**SVL-5**	**SVL-7**	**SVL-9**	**SVL-11**
Fragment range	303-309	282-318	405-452	167-168	265	220	175-176	210-211	135-210	97-98	306-307	105
Number of alleles	3	6	2	2	3	2	2	2	3	1	1	2
Null alleles (%)	78.6	0	0	14.3	78.6	85.7	21.4	35.7	14.3	0	0	92.9
Allelic range	9-10	6-12	8-12	5	4	12	1	2	0--9	1	4	3
Diversity index	38.5	85.7	36.3	26.4	38.5	26.4	36.3	49.5	38.5	0	0	14.2

O121

	R1				R2				R3			
	SVL-1	**SVL-3**	**SVL-4**	**SVL-8**	**SVL-2**	**SVL-6**	**SVL-10**	**SVL-12**	**SVL-5**	**SVL-7**	**SVL-9**	**SVL-11**
Fragment range	0	293-353	418-419	208-209	0	124-305	0	255	169.3-169.5	98	306-315	132
Number of alleles	1	6	2	2	1	6	1	2	1	1	2	2
Null alleles (%)	100	0	0	42	100	7	100	93	0	0	no	93
Allelic range	n/a	6-18	10	8	n/a	2-22	n/a	6	2	1	4-5	7
Diversity index	0	85.7	14.3	49.5	0	60.4	0	14.3	0	0	52.7	14.3

O45

	R1				R2				R3			
	SVL-1	**SVL-3**	**SVL-4**	**SVL-8**	**SVL-2**	**SVL-6**	**SVL-10**	**SVL-12**	**SVL-5**	**SVL-7**	**SVL-9**	**SVL-11**
Fragment range	0	288-305	387-452	167-168	0	0	175-176	210-211	181-197	97-98	306	0
Number of alleles	1	4	3	1	1	1	1	2	3	1	1	1
Null alleles (%)	100	0	14	0	100	100	0	0	0	0	0	100
Allelic range	n/a	6-10	8-12	5	n/a	n/a	1	2	4-7	1	4	n/a
Diversity index	0	57.1	47.3	0	0	0	0	0	58.2	0	0	0

O145

	R1				R2				R3			
	SVL-1	**SVL-3**	**SVL-4**	**SVL-8**	**SVL-2**	**SVL-6**	**SVL-10**	**SVL-12**	**SVL-5**	**SVL-7**	**SVL-9**	**SVL-11**
Fragment range	371	270-317	451-452	155-259	253-277	151-160	278	281-307	169	392	306	0
Number of alleles	2	5	2	3	4	4	4	3	1	3	1	1
Null alleles (%)	93	0	93	51	29	yes	yes	yes	no	yes	no	yes
Allelic range	20	4-12	12	5-12	2-6	5-6	16	9-12	2	38	4	n/a
Diversity index	14.3	50.5	14.3	47.3	62.6	67	39.6	27.5	0	56	0	0

84 non-O157 STEC

	R1				R2				R3			
	SVL-1	**SVL-3**	**SVL-4**	**SVL-8**	**SVL-2**	**SVL-6**	**SVL-10**	**SVL-12**	**SVL-5**	**SVL-7**	**SVL-9**	**SVL-11**
Total fragment range	303-444	270-365	387-452	155-259	253-295	124-305	175-282	210-307	169-210	98-392	306-315	105-160
Total # of alleles	18	13	5	5	9	15	5	6	7	4	2	8
Null alleles	yes	no	yes	yes	yes	yes	yes	yes	yes	yes	no	yes
Allelic range	9-32	4-21	8-12	4-12	2-9	2-23	1-16	2-12	2-9	1-38	4-5	3-14
Total diversity index	59.8	89.1	59	69.5	67.3	75.9	53.1	55.8	53.1	47.6	17.4	32

FIGURES

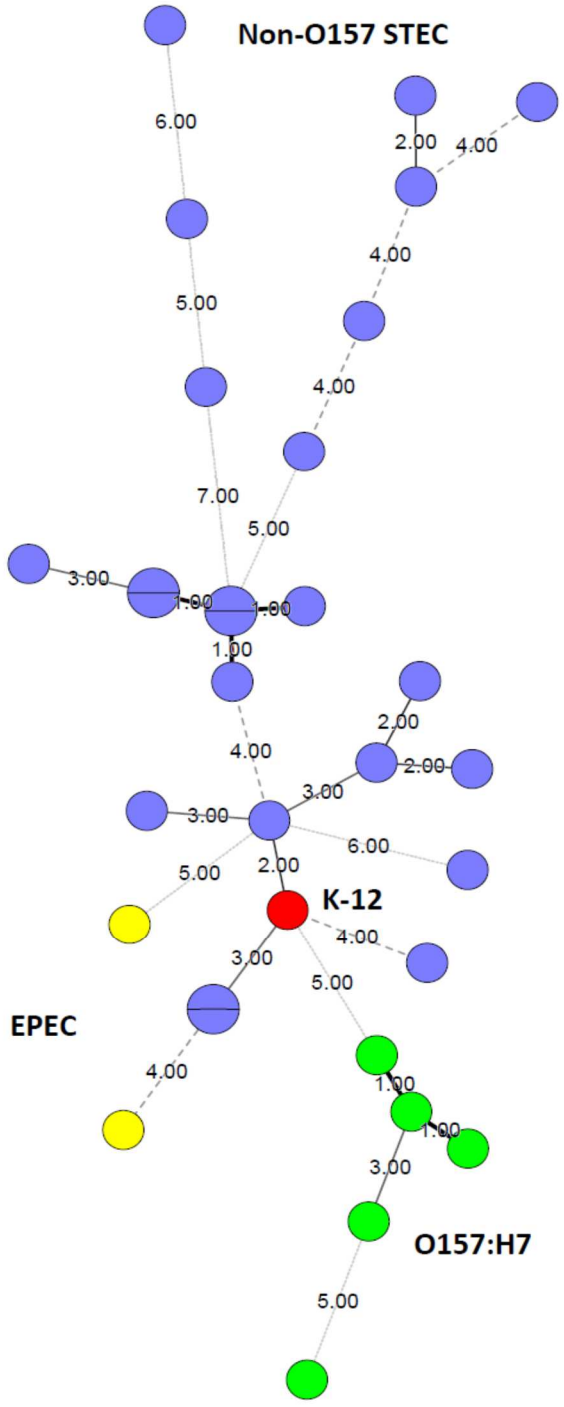

Figure 1. Minimum spanning tree of MLVA types from 24 non-O157 STEC isolates, 5 STEC O157:H7 isolates, 2 EPEC (O119 and O55) isolates, and 1 generic E. coli K-12 isolate. Each circle represents a single MLVA type with the size proportional to the number of isolates with that MLVA type. Numbers on branches indicate the number of loci that differ between each MLVA type.

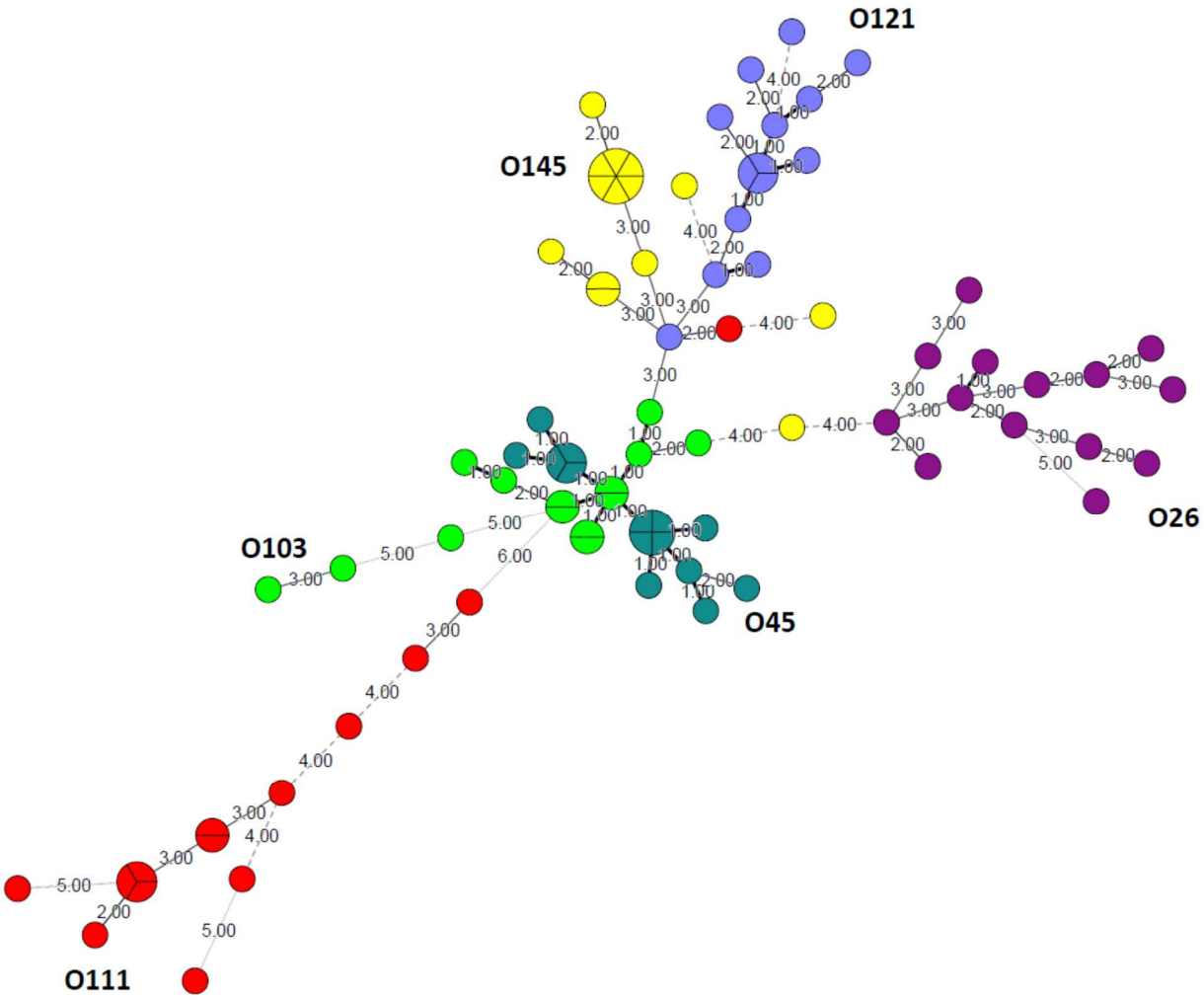

Figure 2. Minimum spanning tree of MLVA types from 84 non-O157 STEC isolates. Each circle represents a single MLVA type with the size proportional to the number of isolates with that MLVA type. Numbers on branches indicate the number of loci that differ between each MLVA type.

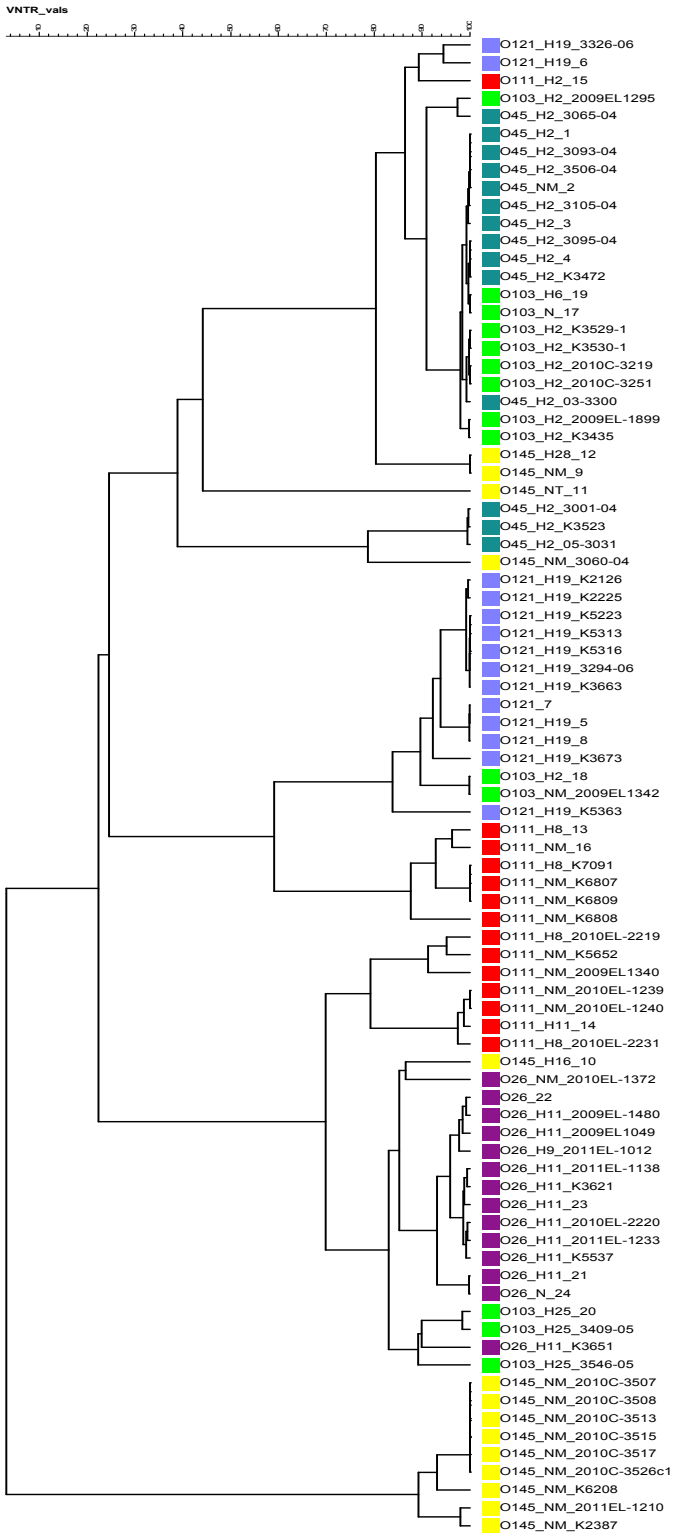

Figure 3. Cluster dendrogram of 84 non-O157 STEC isolates generated by BioNumerics using categorical coefficient and UPGMA clustering.

VITA

Christopher Thomas Timmons

Candidate for the Degree of

Master of Science

Thesis: IMPROVING MOLECULAR DETECTION AND DISCRIMINATION OF *SALMONELLA* AND SHIGA TOXIN-PRODUCING *ESCHERICHIA COLI*

Major Field: Entomology and Plant Pathology

Biographical:

Education:

Completed the requirements for the Master of Science in Entomology and Plant Pathology at Oklahoma State University, Stillwater, Oklahoma in December, 2012.

Completed the requirements for the Bachelor of Science in Biology at the University of Texas at Tyler, Tyler, Texas in May, 2010.

Experience:

Graduate Research Assistant. National Institute for Microbial Forensics and Food and Agricultural Biosecurity (NIMFFAB), Henry Bellmon Research Center, Oklahoma State University, Stillwater, Oklahoma, August 2010 to December 2012.

Undergraduate Research Assistant. Department of Biology, University of Texas at Tyler, Tyler, Texas, May 2008 to August 2010.

Professional Memberships:

American Phytopathological Society
International Society of Food Protection
Entomological Society of America