# GENETIC DIVERSITY OF SHIGA TOXIN-PRODUCING PROPHAGES AMONG *ESCHERICHIA COLI* ISOLATED FROM CATTLE IN MULTIPLE HERDS

By

Yi-An Yang

A THESIS

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

Comparative Medicine and Integrative Biology—Master of Science

2014

ABSTRACT

GENETIC DIVERSITY OF SHIGA TOXIN-PRODUCING PROPHAGES AMONG *ESCHERICHIA COLI* ISOLATED FROM CATTLE IN MULTIPLE HERDS

By

Yi-An Yang

Shiga toxin-producing *Escherichia coli* (STEC) arise after the insertion of a group of bacteriophages encoding the Stx gene. Shiga toxin-producing bacteriophages (Stx phages), notably, are heterogeneous with mosaic genomic structures. However, little is known of the genetic diversity of Stx phages across STEC populations. Here, we examined the diversity of Stx prophages among *E. coli* isolated from 52 animals from 4 Michigan cattle herds between 2011 and 2012 using a novel long range PCR-based restriction fragment length polymorphism (RFLP) method that targets a distinct region (*ciii-stx,* 10-17 kb) within three Stx phage subtypes. Sixteen Stx phage types (PT) were identified; 10 of the 16 Stx PT encoded Stx1, while three encoded Stx2a and another three encoded Stx2c. Interestingly, the same Stx PT was frequently identified within each cattle herd; despite the STEC genotypes are diverse in the herd. These data suggest that transmission of both STEC and Stx phages occurs regularly and provides insight into the diversity and dynamics of specific Stx phages within and across herds. Identifying the distribution of Stx PT present in the cattle reservoir would enable the linkage to STEC epidemiology study, and to further determine whether specific Stx PT contribute to higher virulence in STEC infection.

ACKNOWLEDGEMENTS

First of all, my deepest appreciation goes to my mentor, Dr. Shannon Manning. Without her guidance and insightful suggestions, this project would not have been possible. The thesis is completed under her patient instruction and continuous encouragement. Besides, I would like to thank the rest of my committee, Dr. Steve Bolin, Dr. Matti Kiupel and Dr. Martha Mulks, for their observant comments and feedback.

The current work is achieved by the joint efforts provided by many lab mates involved in the cattle STEC study. Special thanks to Christina Venegas, Akanksha Khare, and Lindsey Ouellette who did the herd sampling and completed the massive STEC isolation process. Advice and comments given by Dr. David Lacher, Scott Henderson and Pallavi Singh have been helpful in LPCR primer design. For all the fellow lab mates who shared the daily life with me working together, those days would not be forgotten. My sincere thanks also go to Enteric Research Investigational Network (ERIN), USDA and MSU Food Safety Group Graduate Assistantship for their financial support. More so, I would like to offer my gratitude to my program advisor, Dr. Vilma Yuzbasiyan-Gurkan, who gave me the warm care over the past two years.

Last but not least, I would like to thank my family, significant other and all the friends, for standing by me spiritually all the time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO SYMBOLS OR ABBREVIATIONS

| | |
|---|---|
| Stx | Shiga toxin |
| Stx phage | Shiga toxin-producing bacteriophage |
| STEC | Shiga toxin-producing *Escherichia coli* |
| LPCR | Long range PCR |
| RFLP | Restriction fragment length polymorphism |
| PT | Phage type |
| NT | Nontypeable serotype |

## Introduction

Bacteriophages, which can insert into bacterial chromosomes, can rapidly change a commensal bacterium into a human pathogen. Shiga toxin (Stx)-producing *E. coli* (STEC), for example, emerged following the acquisition of a Stx temporate lambdoid bacteriphage (Stx phage) and is currently a leading cause of food-borne infections. The clinical outcomes of STEC infection vary from watery diarrhea, hemorrhagic colitis to hemolytic uremic syndrome (HUS) (1, 2). Numerous food-borne outbreaks have been linked to STEC contamination in the U.S. since 1982, and cattle were suggested to represent a natural reservoir with asymptomatic infections (3). Humans are usually infected with STEC through consumption of undercooked ground beef products, raw milk, or contaminated vegetables.

Stx, or verocytotoxin (VT), is the key virulence factor of STEC infection. Following bacterial attachment to the host intestinal mucosa, Stx is absorbed and travels through the blood to the kidneys, where an early local inflammatory event causes occlusion and thrombosis within the glomerular vasculature (4). Stx triggers broad inflammatory response and endothelial cell activation, which leads to secretion of thrombotic factors and induces platelet aggregation (5). Moreover, the globotriaosylceramide (Gb3) cell surface receptor is upregulated during inflammation and promotes Stx internalization in endothelial cells (5). By binding to the 28S ribosomal RNA, Stx acts to inhibit protein synthesis with target cells

(6). In summary, the intense inflammatory response, prothrombotic activation and the direct effect of Stx on the endothelium that promote cell injury result in severe vascular damage and the development of HUS. HUS is characterized by hemolytic anemia, thrombocytopenia and acute renal failure, whereas 20-25% of the patients also develop neurological dysfunction (7).

STEC produces two major Stx subtypes, Stx1 and Stx2, each with multiple variants. Stx1 and Stx2 are genetically and immunologically distinct, and share 56% amino acid identity. Each Stx variant was found to vary in their ability to cause tissue damage (8) and clinical outcomes (9, 10). Stx2, for example, was 1,000 times more potent than Stx1 in human renal microvascular endothelial cells (11), and STEC strains producing Stx2 were found to cause a higher frequency of HUS (10, 12, 13). Moreover, among the Stx2 group, only the Stx2a and Stx2c have been linked to HUS, whereas STEC strains possessing *stx2d* or *stx2e* subtypes have only been isolated from patients with diarrhea or asymptomatic infections (10). Also, several studies demonstrated that STEC O157 strains carrying *stx2a* (with or without *stx2c*) were more frequent among clinical isolates than *stx2c* alone (9, 14), and O157 strains possessing *stx2c* alone are prevalent among bovine isolates and asymptomatic carriers (15).

Notably, the Stx gene is encoded on a group of heterogeneous lambdoid bacteriophages that can integrate into approximately 11 distinct insertion sites in the *E.coli* chromosome (16).

The highly mobile genetic elements can immediately transform the commensal bacteria or the original pathotype into a Stx producer. An important example is the 2011 outbreak in Germany caused by STEC O104:H4, which emerged from an enteroaggregative *E.coli* strain via acquisition of a phage carrying *stx2a* (17). The ability to produce phage-encoded Stx2a combined with the enteroaggregative colonization efficiency resulted in a new highly virulent hybrid strain that contributed to the most severe STEC outbreak in history.

Stx phage is a temperate lambdoid phage with double-stranded DNA and a short tail (*Podoviridae*) or long tail (*Siphoviridae*) structure. The phage genome is characterized by a mosaic structure originating from frequent recombination events between the lambdoid phage family (18). After infecting susceptible bacteria, these phages integrate into the host chromosome and enter a latent prophage state (lysogeny), which possess the ability to form multiple lysogens within the same cell (19, 20). Alternatively, when the prophages detect the host cell damage or other stressors, they can switch to the lytic cycle and start to replicate within the bacteria. Ultimately, the cell is lysed and the free phage particles are released into the environment. Multiple factors have been determined to trigger this event. Environmental stresses such as antibiotics (21-24), DNA damaging agents (25), and UV light (26) as well as host factors including low-iron conditions (27) and reactive oxygen species (28) can induce lambdoid prophages to excise from the host genome and begin reproduction. During the phage replication process, the Stx gene (*stx*) is transcribed and toxin is produced.

The level of Stx production is greatly dependent on Stx phage induction. Previous studies have demonstrated that *stx,* which is located downstream of replication and antitermination genes, is under control of the phage-regulated pR and pR' promoters. Moreover, the amount of Stx1 produced and released are associated with an increase in *stx1* copy number and the regulation of phage-mediated lysis (27, 29). As a result, Stx phages are viewed as the key factors associated with STEC virulence, as they carry the gene encoding different Stx subtypes and control toxin production level. In the literature, the level of *stx* expression and toxin production has been shown to impact the pathogenicity of STEC (30, 31), and the bovine STEC population was found to have more variable Stx production compared to clinical isolates (31, 32). However, one study has found that the Stx genes of STEC isolated from cattle were closely related to those from human-derived STEC (33). These findings present an opportunity to answer the following questions: Does the genetic background or genotype of Stx phages determine the virulence of STEC isolates? Do certain types of Stx phages contribute to enhanced virulence in humans? Do cattle-derived STEC isolates share the same Stx phage types as human-derived isolates?

Studying the diversity of phage genomes has been difficult for several reasons. First, there is no universal gene, analogous to the 16S rRNA gene used for bacteria, exists throughout all phage families. Although family-specific genes, such as the coliphage T4 portal protein g20, has been used as a marker to study the diversity of *Myoviridae* that infect

the marine cyanobacterium *Synechococcu*s (34, 35); this gene is not present in the lambdoid phage family. More recently, Smith et al. developed a multilocus characterization scheme for Stx phages based on individual core functional genes and applied it to 70 *stx*+ lytic phage populations induced from 463 STEC isolates from cattle (36). In their study, *int* (integrase), *n* (early antiterminator), *o* (replication protein), *p* (polymerase), *ci* (repressor protein), *cro* (antirepressor protein), and the genes encoding capsid and host recognition proteins were highly diverse and therefore, multiple PCR-based molecular screening assays were needed to examine the variation. In addition to the extra time needed to isolate Stx phages, the data associated with these assays may not be representative of the phage population given that only the lytic phages could be examined.

The phage isolation-restriction fragment length polymorphism (RFLP) method is considered to be most sensitive and is currently the principle phage typing tool in use (17, 37, 38). This method involves the induction of phages from the *E.coli* host with mitomycin C or other antibiotics and subsequent RFLP digestion of phage DNA following isolation. There are, however, several limitations associated with this method. The induction of prophages among STEC isolates, for example, can result in the propagation of more than one phage originating from the same *E.coli* host. The identification of Stx phages within the lytic phage population requires plaque screening by southern blot hybridization (37) or other time-consuming methods. In addition, some Stx1 phages are non-inducible (37). A prior

study using the phage-isolation RFLP method in 168 STEC strains isolated from cattle of separate herds, for instance, detected markedly different restriction patterns and the infectious Stx phages could only be induced from 30 strains (18%) (39). Therefore, the Stx phage diversity within the same cattle herds could not be accurately measured. Other groups have used phage typing schemes based on bacterial lysis patterns (40, 41), but the results are not always consistent and the restriction patterns of typing phages can be altered following phage propagation, which was shown previously for *Salmonella* (42). In conclusion, the heterogeneous nature of phage genomes and the difficulties associated with Stx phage isolation limit the ability to examine the true genetic diversity of Stx phage populations in epidemiological studies.

To date, only a few number of Stx phage genomes have been fully sequenced and analyzed. Understanding the genetic content of Stx phages allows us to better understand the strategies used by phages to regulate host infection and Stx production. In general, the organization of ~80 genes in the Stx phage genome can be divided into two parts, the functional genes in the early region and the structural genes in the late region (43). The functional motifs within the early region are generally in the following order: integration, regulation circuit, phage immunity, replication, Stx production and host cell lysis. The late gene region, however, encodes a cluster of structural proteins involved in DNA packaging and the genes for head and tail assembly. The lambdoid phage family employs the CI-Cro

epigenetic switch mechanism to control lysogeny and entry into the lytic cycle, as well as phage immunity to prevent infection by new phages that may be homologous to the existing prophage within host chromosome (44). Three proteins are involved in the establishment and maintenance of lambdoid phage repression - CI, CII and CIII. Among them, CI is the main "λ-repressor" that acts by repressing RNA synthesis from early genes for replication and entry into the lytic cycle. CII is the transcriptional activator for the initial synthesis of the CI repressor. CIII, lastly, serves as a protease inhibitor to enhance the stability of CII and the lysogenic response (45). In contrast, Cro prevents synthesis of new CI protein hence impedes the stable lysogeny state (46).

Despite the universal genetic organization of Stx phage genomes, individual genes within these genomes are heterogeneous with numerous variants. Indeed, being the most abundant organisms in the biosphere, each phage is subjected to tremendous selective pressures to outcompete others. The extensive level of horizontal gene exchange between phages has resulted in the characteristic patchy, mosaic structure of the Stx phage family, which has led to an increase in the number and diversity of Stx phage types present. Six characteristic regions with remarkably different nucleotide sequences have been identified within the phage genome across Stx phages. Among them, region V and VI were considered to be the most distinct portion and are important for controlling phage immunity, replication, and host cell lysis (47). Extreme variation was specifically found in the genes

encoding the integrase and capsid as well as genes associated with the genetic switch between lysogeny and lytic cycle (*ci*, *cro*) that drive the formation of novel phages and extend the bacterial host range (18, 37, 39, 48). Also, there are many morons (49), or DNA elements inserted between genes in prophage genomes, which are not important for the phage life cycle (50), as well as hypothetical proteins (43) with unknown functions. As a result, only a few genes are conserved among all Stx phage genomes. These conserved genes include *q* (late antiterminator), *cii* (transcriptional activator of CI), *ciii* (protease inhibitor for CII), *stx* (Shiga toxin), *roi* (DNA binding protein) as well as the gene encoding the tail spike protein in short-tailed phages (36, 43, 51). The antiterminator Q gene, which controls the expression of late phage genes, is located upstream of s*tx* and represents the most conserved motif across the Stx phages (52). Nonetheless, certain Stx phages were found to completely lack some of these conserved genes, particularly those encoding proteins CII and CIII (17).

The goals of this study are to: (1) develop a more sensitive Stx phage genotyping tool; (2) examine the genetic diversity of Stx bacteriophage in STEC isolates from multiple cattle herds and deer in Michigan; and (3) compare the Stx phage type (PT) in a cattle- and human-derived STEC isolates. We hypothesize that Stx phage transmission occurs within animals from the same geographic location. Evidence for this hypothesis is that the Stx temporate phages can undergo both the lytic and lysogeny cycle, and frequent gain and loss of Stx-phages across STEC isolates has been documented in the literature (53-55).

Moreover, the dissemination of similar Stx phage types has been observed in the same cattle herd (37) or over a long period of time (17) and studies have detected a high prevalence of Stx phages in urban sewage and animal wastewater (56).

In an effort to better understand the genetic diversity and transmission dynamics of Stx phages in cattle, a major source of human STEC infection, we developed a long range PCR-based RFLP (LPCR-RFLP) phage typing scheme that is capable of characterizing the Stx phage genome while it is integrated into the STEC chromosome. The application of LPCR-RFLP was first conducted to genotype *Haemophilus influenzae* isolates in 1995 (57). The methodology has also been applied to multiple bacteria species and shown to agree with pulsed-field gel electrophoresis (PFGE) data (58, 59) and cluster analyses of published bacterial genomes (60). One key advantage of this technique is that the LPCR primers can be designed to target Stx phage-specific genes without the need to induce and isolate the entire phage population. Here, we demonstrate that one LPCR assay targeting the Vd (*ciii-stx*, 10-17 kb) region of three Stx phage subtypes can adequately differentiate the Stx phage population among 87 STEC isolates recovered from cattle and deer. By this means, we were able to examine the genetic variation in Stx prophages across cattle from different herds, and measure Stx phage transmission dynamics among cattle and deer sharing the same pasture. An exploration of the types of Stx prophages circulating in cattle-derived STEC isolates can enhance our understanding of how these mobile genetic elements move in nature

and serve to drive the evolution of commensal *E. coli* into virulent toxin producers.

## Materials and methods

### *Sample collection and STEC isolation*

In 2011-2012, four herds (three beef, one dairy) were examined as part of a larger study to identify risk factors for STEC shedding in cattle. Fecal grab samples were collected from 312 healthy cattle. Among them, 24 animals were followed at two week intervals for a total of four visits. In addition, a separate study was performed in 2012 to investigate STEC shedding in wild deer that shared the same pasture with one dairy herd. Fecal pellets collected from 34 deer were verified by animal genotyping and subjected to STEC culture. Five grams of feces was homogenized and enriched in *E. coli* broth within 24 hours of collection. Following overnight growth at 37$^{o}$C, bacteria culture were plated onto STEC CHROMagar (CHROMagar, Paris, France) and sorbitol MacConkey agar for incubation overnight at 37$^{o}$C. To further select for the O157 serogroup, an immunomagnetic separation (IMS) assay of *E. coli* O157:H7 using Dynabeads® MAX *E. coli* O157 (Invitrogen Corporation, California, USA) was performed. After IMS selection, the beads were plated on STEC CHROMagar for overnight culture. Up to 20 suspected STEC colonies were recovered per plate and subjected to DNA isolation and PCR confirmation. The herd prevalence of STEC positive animals ranged from 9.6 % to 50% (Table 1).

***Characterization of bacterial strains***

DNA from suspect isolates was examined for the genes encoding Stx1, Stx2 (*stx1*, *stx2*) and intimin (*eaeA*) by KAPA 2G Fast multiplex PCR (Kapa Biosystems, Boston, Massachusetts) using primers from several prior studies (Table 2). The *stx2* positive strains were further differentiated by digesting with *Fok*I to distinguish between *stx2a* and *stx2c* variants using a published method (61) and *stx2a*, *stx2c* specific primers. The determination of O antigen type (serotype) was performed using two separate multiplex PCR assays with primers targeting O157 and six non-O157 serotypes (O26, O45, O103, O111, O121, O145) as well as an internal control encoding malate dehydrogenase (*mdh*). Isolates that failed to amplify one of the seven O-types were considered nontypeable (NT). A subset of 45 STEC isolates representing multiple serotypes were sent to the Food and Drug Administration (FDA) for serotyping via a SNP-based array platform (unpublished protocol)..

Among the subset of 174 STEC isolates recovered from four herds, three to seven unique virulence gene profiles and serotypes were identified per herd. When isolates from the same animal exhibited similar virulence gene profiles and serotypes, only one was selected for inclusion in the phage typing study leaving 82 distinct isolates from 52 cattle for characterization. A subset (n=34) of these isolates came from the same animal at two to four different sampling times that were roughly two weeks apart (Table 1).

For comparative purposes, a subset of eight STEC isolates recovered from deer fecal

pellets located near the dairy pasture was included to determine whether Stx phages were shared across species. Similarly, a subset of 22 strains from humans with STEC O157 infections in Michigan, previously characterized by single nucleotide polymorphism (SNP) genotyping (61) and representing the predominant phylogenetic clades, was also evaluated

***In silico analyses of Stx-bacteriophage genomes:***

Eighteen full-length Stx phage genome sequences (Table 4) were selected from the National Center for Biotechnology Information (NCBI) database as reference strains. Multiple homologous conserved genes previously identified in comparative genomics studies (36, 43) and the PhiSigns phage database (62) were retrieved from the Stx phage reference strains for comparison. Alignment was implemented in MegAlign (ver. 5.06) using the ClustalW algorithm. Five sets of LPCR primers were designed within or spanning region V and VI, which were found to be the most distinct regions of the entire Stx phage genome (47). Primers for *q*, *ninG, roi, ciii, stx* and *j* encoding the Q antiterminator protein, NinG recombination protein, Roi DNA binding protein, CIII regulatory protein, Shiga toxin and tail fiber protein, respectively, were designed based on the universal sequence identity (Table 3). In each primer set, either forward or reverse primers were designed to target one of the three *stx* variants (*stx1, stx2a, stx2c*) to ensure the amplification of each Stx phage subtype, which is dependent on the *stx* gene profiling results. The five LPCR regions corresponded to the

following regions: *q-stx*, *ninG-stx*, *roi-stx*, *ciii-stx*, *stx-j*, and were designated region Va, Vb, Vc, Vd and VIa, respectively (Fig. 2). The forward primer ciiiV2-F targeting *ciii*, which encodes the CIII regulatory protein, was obtained from Shima et al (58) and paired with reverse primers targeting s*tx2a* or *stx2c*, depending on Stx2 subtyping results (Table 3). For amplification of *ciii* in the Stx1 phages, a distinct primer was used as there is a SNP within *ciii* in the Sakai Stx1 phage genome that would prevent amplification with the Stx2 phage *ciii* primer.

The 18 Stx phage reference strains were trimmed into the regions Va, Vb, Vc, Vd and VIa flanked by two primers. The LPCR regions were digested *in silico* using the electronic gel function of NEBCutter V2.0 (New England Biolabs, Inc.) to compare banding patterns for 10 common restriction enzymes (*Bc1*I, *Bgl*I, *BstU*I, *CviAI*I, *EcoR*V, *Fat*I, *Fsp*I, *HinP1*I, *Msp*I, *Pst*I) that generate 5 to 15 fragments. Restriction enzymes were chosen based on the ability to differentiate RFLP band patterns for each of the Stx phage reference strains.

***PCR-based RFLP Phage genotyping assays:***

Amplification of Va, Vb, and Vc were conducted using LPCR with the KAPA 2G Fast multiplex PCR kit. The reaction began with denaturation for 3 min at 95°C, followed by 30 cycles of replication (15-s at 94°C, 30-s at 64°C and 30-s at 72°C) and final extension at 72°C for 3 min by using a GeneAmp PCR System 9700 (Folster city, CA). Positive LPCR

products were digested for two additional hours with 5 U of *Msp*I or *HinP1*I and analyzed by 3% agarose gel electrophoresis in 0.5X TAE buffer for 4-5 h followed by ethidium bromide staining.

LPCR for regions Vd and VIa were performed as described by Shima K et al (58) using the long and accurate (LA)-PCR Kit (Takara Bio Inc., Otsu, Japan) with DNA templates adjusted to 50-150 ng/ul. The reactions began with denaturation for 1 min at 94°C followed by 30 cycles of: 20 s at 98°C and 10 min at 68°C with a final extension at 72°C for 10 min. These assays were performed using the Bio-Rad iQ5™ PCR System and the melting temperature for Stx2c phage was adjusted to 66°C for 5 min. Positive LPCR products with single bands were measured by Qubit® 2.0 Fluorometer (Invitrogen™, Carlsbad, CA) to estimate the DNA concentration. The DNA concentration of each LPCR amplicon was normalized to 30 ng/ul prior to restriction digestion and RFLP analyses. The amplicons were digested for 2 h by 10U of *Bgl*I and 8U of *Eco*RV and then analyzed by 2% agarose gel electrophoresis in 1X TAE buffer for 4-6 h. A 1-kb plus DNA ladder (Invitrogen) was used as the molecular mass standard. Gel photographs were taken using the Gel-Doc XR$^+$ system (Bio-Rad, Hercules, CA) and the band patterns were analyzed and grouped using BioNumerics software (ver. 5.1; Applied Maths, Austin, Texas). The dendrogram was created based on the UPGMA cluster analysis of dice index of normalized RFLP patterns. Each Stx phage type characterized from cattle-derived STEC isolates were designated by 'P'

in front, followed by a single number, whereas each distinct Stx phage type identified from clinical STEC strains were designated by letter "Φ" before the number. A flow chart for current study design is showed in figure 1.

***Repetitive element PCR (Rep-PCR) and stx sequencing***

Rep-PCR was performed using BOX primer to generate DNA fingerprints from STEC isolates as described in previous study (63). The DNA templates were adjusted to 50 ng/ul before amplification and the products were analyzed by 2% agarose gel electrophoresis in 0.5X TAE, followed by ethidium bromide staining. The Rep-PCR band patterns were analyzed using BioNumerics software (ver. 5.1; Applied Maths, Austin, Texas).

STEC isolates were sent for sequencing as part of a larger study using Illumina next generation sequencing tool. The Stx genes were extracted from 19 strains in one dairy cattle herd and a subset of 3 deer sharing the same pasture. Sequences were analyzed in MegAlign (ver. 5.06) using the ClustalW algorithm for alignment.

## Results

### *In silico analysis of Stx phage genomes*

Regions Va, Vb, Vc, Vd and VIa within the STEC O157 Sakai strain Stx2a phage (Fig. 2) genome generated 1.7 kb, 2 kb, 3 kb, 12 kb and 16 kb products, respectively. The amplicon sizes, however, varied considerably among the Stx phage reference strains (Table 4). Restriction patterns of regions Va (*q-stx*), Vb (*ninG-stx*) and Vc (*roi-stx*) using *Msp*I and *HinP1*I enzymes could only categorize the 18 reference strains into five groups. All three assays were therefore unable to adequately differentiate genetically distinct Stx phages

Region VIa (*stx-j*), located downstream of *stx* and spans the VI region identified by Sato et al. (47), was evaluated next. *In silico* analysis of the LPCR targeting region VIa generated 16 kb to 30 kb products in the reference strains (Table 4). Notably, the restriction band patterns using *Pst*I, *Fsp*I enzymes for Stx2a phages and *EcoR*I, *Bcl*I enzymes for Stx1 and Stx2c phages were capable of classifying the 18 Stx phage reference strains into 17 groups (Fig. 3). Identical restriction patterns were found in Bacteriophage 933V and Stx2 converting phage I. Hence, region VIa appears to more adequately differentiate distinct Stx phages when compared to the three assays targeting Va, Vb and Vc. Alignment of region VIa in each reference strain revealed that the RFLP patterns can accurately differentiate Stx phages types.

The final region evaluated in this study is region Vd (*ciii-stx*), which results in 10 kb to

17 kb LPCR products among the reference strains (Table 4). Although region Vd adequately differentiated 10 of the 13 reference strains *in silico*, several phage genomes could not be examined due to the absence of *ciii* or insertion mutations in *stx*. These Stx phage reference strains include: bacteriophage 1639 (AJ304858), enterobacteria phage VT2phi_272 (HQ424691), *Escherichia* phage TL-2011c (NC019442), *Escherichia* phage P13374 (HE664024) and Stx2 converting phage vB_EcoP_24B (HM208303). In all, the restriction band patterns of region Vd using *EcoR*V and *Bgl*I enzymes generated three to eleven bands, though identical restriction patterns were observed for Bacteriophage 933V and VT1 phage Sakai, Bacteriophage 933W and Stx2 converting phage I, VT2 phage Sakai and Stx2 converting phage II. For confirmation, the nucleotide sequences of region Vd from the 13 reference strains were aligned and showed 99-100% sequence similarity between strains with the same RFLP profile. Importantly, the RFLP patterns can differentiate Stx phages types at 93.9% sequence similarity (Stx1 converting phage and Stx2 converting phage II, see Fig. 5).

***LPCR-RFLP profiling of Stx phages in STEC control strains***

To validate the applicability of the LPCR-RFLP Stx phage typing method, each assay targeting the different Stx phage regions was evaluated using *stx1*/*stx2*-positive STEC O157:H7 control strain TW08264 and s*tx2*/*stx2c*-positive STEC O157:H7 control strain TW14359 from the Sakai and Spinach outbreaks, respectively, and a subset of ten

cattle-derived STEC isolates from different herds. The cattle-derived STEC isolates were selected based on the assumption that Stx phages would be distinct in strains from different herds and similar in strains from the same herd. Regions Va, Vb and Vc were successfully amplified with expected product sizes and restriction patterns as the *in silico* analyses. However, identical RFLP band patterns were found in cattle across three herds as well as the deer and four randomly-selected calf-derived STEC isolates acquired from the Diagnostic Center for Population and Animal Health. Consequently, all three assays failed to adequately differentiate the Stx phage types between herds and animals and should not be used for epidemiologic and phage diversity studies.

Regions Vd and VIa were also examined in the two STEC O157:H7 control strains and two of the 10 cattle-derived STEC isolates. Both regions were successfully amplified in each strain and a single product of 12 kb and 16 kb was generated for regions Vd and VIa, respectively. Restriction digestion with *EcoR*V and *Bgl*I enzymes in region Vd detected a distinct RFLP pattern for each strain that matched the predicted fragment sizes in the *in silico* analysis. Although the LPCR products of region VIa were relatively faint, the RFLP banding patterns were in concordance with the *in silico* profiles. Amplification and digestion of both regions VIa and Vd yielded distinct banding patterns in four strains using each assay, suggesting that both regions can adequately differentiate Stx phage types (Fig. 9). Because Sato et al. (47) had previously determined that region Vd was the most diverse

across Stx phages and the LPCR assay had better resolution, it was examined in a larger set of STEC isolates to better understand phage diversity and distributions.

***Stx phage diversity and distributions in STEC from multiple cattle herds and deer***

A total of 96 unique STEC strains originating from 62 cattle in four herds and four deer were evaluated by LPCR-RFLP. However, 14 strains from 10 animals cannot generate the LPCR products using current primer set. Restriction digestion with *EcoR*V and *Bgl*I enzymes yielded successful LPCR products with distinct banding patterns that could be differentiated visually. Among all 82 STEC isolates from cattle, 34 isolates were from different visits (Table 7). A total of 15 distinct Stx phage types were identified overall and between two and 14 phage types were recovered per visit with no more than five types per herd at any one time point (Fig. 7).

The distribution of Stx 1, 2a and 2c phage types also varied among the 82 unique STEC strains from the four cattle herds (Table 5). For the 45 *stx1*-positive STEC strains, a total of ten different phage types were detected from three herds. For the 14 *stx2a*-positive strains, three different phage types were recovered in two herds, whereas three phage types were found in 42 *stx2c*-positive strains in two herds. Herds with a greater number of *stx* profiles, for instance, had more phage types though one type typically dominated in frequency per herd. For example, Stx1 phage type (PT) 2 and 3 and Stx2c PT4 were recovered from most

(12%, 24%, 70%) of the unique STEC isolates in herd 8, while the Stx2c PT12 predominated in 85.7% of unique isolates from herd 11. For herd 12, however, the Stx1 PT5 (71.4%) and the Stx2a PTs 10 (28.6%) and 15 (50%) were most common. By contrast, the four *stx1*-positive STEC isolates from four deer were less diverse with only two phage types. Moreover, among the few animals with more than one unique strain, eight of the 11 shared the same phage types between isolates, whereas seven animals carried multiple unique strains with different phage type combinations at any visit, but each phage type was shared with other animals in the same herd.

During the first visit, the greatest number of unique STEC strains was isolated from herd 8 where only three specific phage types (PT1, PT3, PT4) were identified in 14 cattle (Table 6). Among them, Stx2c PT4 was recovered from 85.7% of animals, whereas Stx1 PT1 and Stx1 PT3 were each found in one animal. In herd 9, five distinct phage types (PT5, PT6, PT7, PT8, PT10) were identified in nine cattle, but only Stx1 PT5 (22.2%) and Stx1 PT6 (55.6 %) were recovered from more than one animal. Notably, Stx1 PT6 was also identified in two deer and the other two carried Stx1 PT11. In herd 11 (n=4 animals), only two phage types (PT12, PT13) were recovered; Stx2c PT12 was found in three O157 strains from three animals and Stx2c PT13 was found in a NT strain from the remaining animal. In herd 12, five distinct phage types (PT5, PT10, PT14, PT15, P16) were identified in 14 animals carrying unique STEC strains, and were present in 71.4%, 28.6%, 7.1%, 50% and 14.2% of

animals, respectively. Notably, most of the unique STEC strains tested carried both Stx1 PT5 and either Stx2a PT10 (28.6%) or Stx2a PT15 (42.9%) in the same O157 strain, while Stx1 PT14 and Stx2a PT16 were present in three NT strains separately. Overall, each cattle herd had between two and six phage types, which were driven by the distribution of *stx* profiles.

A change of Stx1 phage type over time was observed in herd 8 (Table 7). For example, Stx1 PT2 was recovered from four animals at visit II, but starting from visit III to IV, all nine animals carrying *stx1* shifted to Stx1 PT3 instead. Moreover, animals 6, 10 and 15 carried *stx1*-positive STEC isolates of serotype O103:H2 or NT at visit II and O26:H11 at visit III, but the Stx1 phage type changed from P2 to P3.

***Inter- and intra-species transmission of Stx phages***

Within each of the cattle herds, the range of animals sharing the same phage type as another animal at the same visit is between 14.3% and 85.7%, which varied across each of the four visits. For visit 1, 39 of the 45 animals examined in all 4 herds shared the same phage type with at least one other animal in the same herd, which represented eight distinct Stx phage types (Stx2c PT4, Stx1 PT5, Stx1 PT6, Stx2a PT10, Stx1 PT11, Stx2c PT12, Stx2a PT15, Stx2a PT16) (Table 6). At visit 2, four and seven animals in herd 8 shared Stx1 PT2 and Stx2c PT4, respectively. At visit 3, unique STEC strains were recovered

from herds 8 and 11, where seven and four animals shared Stx1 PT3 and Stx2c PT4 in herd 8, and two animals shared Stx2c PT12 in herd 11. At visit 3, two and five animals shared Stx1 PT3 and Stx2c PT4 in herd 8, and Stx2c PT12 was recovered again in one animal in herd 11 (Table 7). Herd 8 was interesting in that 19 of the 22 STEC positive animals had Stx1 PT4. Only eight of these 35 strains, however, represented serotype O6:H34, while the remaining 25 were nontypeable (NT) and two strains were identified as serotype O157:H7 and O103, suggesting that the same phage types were circulating among different STEC strain types.

In addition to phage transmission within animals of the same herd at one visit, four of the total phage types (PT2, PT3, PT4, PT12) have been observed in more than one visit. In herd 8, for example, Stx2c PT4 was recovered throughout the four different sampling times in 28 STEC isolates. Evidence of Stx phage transmission was also observed between different *E.coli* hosts as the same Stx phage type found in different STEC serotypes. For instance, an O103 STEC isolate from animal 5 possessed both Stx1 PT2 and Stx2c PT4 at visit II, and an O157:H7 STEC isolate from animal 11 was found to carry both Stx1 PT3 and Stx2c PT4 at visit III. Similarly, most (71.4%) of the 35 strains with Stx2c PT4 were NT, though eight strains were identified as O6:H34 by the FDA array typing platform. Additionally, we also observed the emergence of phage types over time due to the influx of novel STEC strains carrying distinct phage populations. At visit II, for example, four out of six strains with

Stx1 PT2 were O103:H2 STEC, whereas ten out of 11 strains with Stx1 PT3 were O26:H11 STEC strains at visit III and IV. Since PT3 had not been detected in these animals previously, these data suggest a change in the STEC population occurred over time rather than phage transmission.

When examining the phage type distribution across herds, only two phage types could be identified in more than one herd. Specifically, Stx1 PT5 was identified in two animals (18.2%) from herd 9 and ten animals (71.4%) from herd 12; whereas Stx2a PT10 was found in one animal from herd 9 (9.1%) and four animals (28.6%) from herd 12. It is worth emphasizing that all Stx1 PT5 and Stx2a PT10 were found in O157 strains, which had identical restriction patterns as the Stx1 and Stx2a phages in the STEC O157:H7 Sakai outbreak strain.

Finally, there was also evidence of transmission between cattle and deer belonging to herd 9. An identical Stx1 PT6 was recovered from two deer and five cattle from dairy herd 9, which shared the same pasture during sampling (Table. 6). This suggests possible Stx phage or STEC transmission between cattle and deer living in close geographic locations. Based on the Rep-PCR genotyping profile of the STEC isolates, identical STEC genotypes were only observed in one STEC isolate from dairy herd 9 (animal 32) and one deer (animal 33); both isolates shared the same Stx1 PT6, suggesting STEC transmission versus PT6 transmission. Nonetheless, other STEC isolates from dairy herd 9 carrying the same Stx1

PT6 (animal 25, 27, 28, 30) were found to have diverse STEC Rep-PCR patterns (data not shown). This finding further supports the evidence of Stx phage transmission across different *E. coli* hosts.

***Identification of clinical Stx phage types***

We further employed the Stx phage LPCR-RFLP assay in region Vd to identify the Stx phage type in a subset of clinical O157 STEC isolates from clades 2, 7 and 8 (Table 8). For the most part, distinct population of Stx phages were identified in the clinical STEC isolates. Eleven unique Stx phage types (Φ1-Φ11) were found in addition to the 16 Stx phage types from cattle and deer-derived STEC isolates. Interestingly, specific phage types, which shared the same RFLP patterns with the Stx1, 2a or 2c phages in the STEC O157:H7 Sakai and Spinach outbreak strains, were frequently identified in the clinical isolates. For example, both Stx1 PT5 and Stx2a PT10 were identified in 66.7% and 12.8% of human-derived STEC O157 strains belonging to clades 2 and 8, respectively, whereas Stx2c PTΦ9 associated with Spinach outbreak was recovered from 20% and 71.4% of STEC strains belonging to clades 7 and 8. Moreover, few distinct Stx phage types were recovered from specific clades with higher frequency than others. For instance, Stx1 PTΦ7 (40%) and Stx2c PTΦ11 (40%) were only identified in clade 7 strains, whereas Stx2a PTΦ3 (42.9%) was only recovered in clade 8 strains.

***stx sequencing***

The nucleotide sequences of the Stx1 and Stx2a genes in isolate 791-1 were identical to the Stx gene in the STEC O157:H7 Sakai strain, which also showed the same Stx phage type based on LPCR-RFLP band profile. The Stx2c gene in one STEC isolate from animal 873 exhibited 100% homology to previously published STEC isolates from cattle in Poland (O183:H18 strain CB12500, FR850033). However, the *stx2c*$^{+}$ STEC isolates could not be recognized by our Stx phage LPCR primer set. For the Stx1 gene, 14 STEC isolates from cattle in herd 9 possessed an identical *stx1* with STEC isolates from three deer and reference strains CP-1639 and YYZ-2008. However, despite sharing the same Stx1 gene, the Stx phage LPCR-RFLP band pattern from those animals were not identical to Stx1 phage CP-1639 and YYZ-2008 as compared with *in silico* result, thereby providing support for the ability of our phage typing method to accurately differentiate Stx phages across distinct STEC strains.

**Discussion:**

Stx production, which is encoded by genes carried on Stx phages, is the key virulence factor in STEC infections. Few studies, however, have examined the diversity of Stx phages in epidemiological scale. In particular, little is known about how variable the different Stx prophages are across STEC strains and whether genetic variation in Stx phages is linked to disease severity or transmissibility in reservoir species like cattle. Difficulties associated with phage isolation and characterizing Stx prophages while they are incorporated into the STEC genome has hampered these efforts. In this study, we have shown that a novel LPCR-RFLP method targeting a distinct region Vd (*ciii-stx,* 10-17 kb) of three Stx phage subtypes could be used to examine transmission dynamics and Stx prophage diversity among 86 unique STEC isolates recovered from cattle. We also demonstrated that there is considerable variation in Stx phage populations across cattle from different herds over time and that most cattle within a herd shared the same phage type.

For developing a Stx phage typing scheme, we tested five different regions within the three Stx phage genomes to elucidate the genetic diversity of phage across herds. Among the long regions screened in the present study, primers targeting region Va (*q-stx,* ~1.7 kb) could amplify the Stx phage in all strains, suggesting a universal presence. After digestion with *Msp*I, however, most of the cattle-derived STEC control strains had identical banding patterns and the *in silico* analysis among the reference strains could only distinguish five groups. This result indicates that this region is highly conserved across Stx phages;

therefore, it may be an important target for the development of diagnostic tools aimed at detecting STEC in clinical samples and food matrices rather than evaluating Stx phage diversity. Indeed, several researchers have developed assays to quantify Stx phages from environmental samples by targeting the Q and Stx genes located within this region (52). Other researchers have also found a high degree of similarity in the *q-stx* region among 49 STEC clinical isolates from humans in Germany (64), though an evaluation of cattle-derived STEC isolates has not been done.

Because region V (*ciii-ninG*, 8-12 kb) was found to be the most distinct portion in the entire phage genome in two previous studies (47, 65), we extended this region (Vd, Fig. 2) by changing the reverse primer targeting the Stx gene (*ciii-stx*, 10-17 kb). Moreover, utilizing different primers targeting each *stx* variant (*stx1*, *stx2a*, *stx2c*) can provide the advantage of amplifying Stx phage subtypes separately during co-infection of the bacterial host, thus increasing the specificity. According to Sato et al. (47), region V is the most distinct region and shows remarkable differences between Stx2 phages isolated from different STEC O157 outbreaks. Region Vd also contains the genes that control immunity (*ci*, *cro*), replication (*o*, *p*) and early antitermination (*n*), which were found to be highly diverse with multiple variants in previous study (36). One hypothesis for a high frequency of polymorphisms in this functional gene region is that it may provide an advantage for phages to outcompete other lamdoid phages and successfully infect the bacterial host. In addition, region Vd yielded the

most clear banding patterns in our assay and therefore, it was considered the preferable target for our epidemiological study.

Region VIa spanning *stx* to *j*, which encodes the tail fiber gene, was also found to be a heterogeneous region by *in silico* anaylsis of 18 reference strains. The capsid gene included in this region was previously described to be divergent (36), as well as the R and S genes responsible for host cell lysis (47). The gene encoding the tail spike protein is well conserved across short-tail Stx phages and thus is a desirable primer target (36, 51). However, the large size of the DNA fragment resulted in low amplification frequencies and less amount of amplicons for restriction digestion. Moreover, the tail fiber gene in Stx1 and Stx2c phages was universal among the prophages present in the *E.coli* chromosome with approximately five to ten copies of loci, which prevented from successful long amplification. We utilized this region downstream of *stx* in a subset of Stx2a phages from herds 9 and 12 to confirm the results generated from region Vd. As expected, no distinct RFLP patterns were observed in the VIa downstream region when the RFLP patterns in upstream region Vd were identical (Fig. 9), suggesting that both assays can adequately differentiate phage types. Even though a number of Stx phage reference strains could not be amplified because of a missing CIII gene (17, 43, 66, 67), a fragment could be amplified for most (86%) of the animal-derived STEC isolates examined in the study. Collectively, these data provide additional support for the sensitivity of our Stx phage genotyping assay targeting region Vd

as the combined LPCR regions (~30 kb) comprise over half of the Stx phage genome (50-66 kb) and represent the most divergent areas.

One limitation of this LPCR-RFLP assay is that we are only screening the Stx prophage population that remains integrated into the STEC chromosome, which may not represent the true diversity of Stx phages found in the environment. Also, we might possibly select for certain types of Stx phages using the current primer set, as the CIII gene was found to be absent in four Stx phage reference genomes and *stx* variation has been described (17, 43). Another limitation is that the DNA fingerprint generated by RFLP digestion cannot detect single nucleotide changes within the phage genome. Consequently, future metagenomic studies are still needed to unravel the diverse picture of phage populations present in the environment. Nonetheless, our method was still sensitive enough to allow us to examine diversity and transmission dynamics of Stx prophages as well as STEC strains in important reservoir species.

Overall, we identified 15 distinct phage types among 82 unique STEC isolates recovered from cattle, suggesting that the phage diversity is over five times lower than STEC diversity. This is in contrast to previous finding that variable Stx2 phages were induced from six highly related clinical *E.coli* O157:H7 isolates with identical PFGE pattern (68). Because Stx phages could not be amplified from 14 strains, it is possible that we have underestimated the diversity of this phage population. Since our assay examines the Stx prophage population

while it is integrated into the STEC chromosome, it is possible that an additional diverse group of Stx phages may have entered the lytic cycle during subculture and thus, could not be characterized. As documented in the literature, temperate phages are more prone to enter the lytic cycle under stressful conditions (23, 25, 26, 28). Indeed, we observed loss of *stx* in multiple STEC isolates after subculture. However, when using the current LPCR-RFLP method in original isolates, the Stx gene can be recovered but often revealed a distinct Stx phage type (eg, P1, P7, P8) other than those frequently present in the same herd. It is therefore possible that only a subset of Stx phages remain in the lysogeny state, which could be due to mutations in genes important for excision, for instance. Consequently, this small subset of phages could be preferentially targeted via our LPCR assay, which would result in a lower level of phage diversity. Regardless of this limitation, we were still able to identify multiple phage types and importantly, better understand transmission dynamics of Stx phages in cattle.

In the present study, we observed frequent sharing of identical Stx phage types between animals in the same herd or pasture, and even between STEC hosts with different genetic characteristics. Another study, for example, found that identical Stx phage types were isolated from a limited number of animals in the same cattle herd (37). Moreover, similar Stx phages have been identified from some STEC strains over long periods of time (17) as well as in different geographic locations (47, 64). For example, the genome of

bacteriophage 933V isolated from STEC O157:H7 EDL933 outbreak strain in US in 1982 is identical to the VT1 phage Sakai originated from STEC O157:H7 Sakai outbreak strain in 1996 in Japan. The reason behind these observations could be due to enhanced dissemination of certain Stx phages through commercial products and foods or to the hardiness, mutation rate or abundance (56) of specific bacteriophages in the environment.

To our knowledge, this study is the first to uncover Stx prophage diversity within and between multiple cattle herds and cattle over time. Different Stx phage genotypes may play a key role in determining the virulence of different STEC strains (69). It has been shown that the human and cattle *E.coli* O157:H7 isolates often cluster in separate lineage (70), which are in concordance with our finding that different groups of Stx phage types represented in bovine and clinical STEC strains, although a few numbers of Stx phage types were shared in these two populations. Recently, Park D et al.(71) demonstrated defective Stx2a and Stx2c prophages found in bovine STEC isolates, with either gene deletions or insertional inactivation by IS*1203* in the toxin gene, had an impaired ability to produce Stx. The defect within Stx phage genomic content might be a consequence of adaptation to the bovine host, which also reflects the loss of virulence characteristics necessary to cause disease in human. In contrast, specific Stx2 phage types have also found to correlate with STEC outbreaks and high level shedders in cattle (17, 41, 72). Due to the extensive heterogeneity and frequent mobility of the temporate phage population, the impact of

different Stx phage types on clinical infection is still unknown. Indeed, we found that the human-derived phage types examined in this study were more diverse than in cattle STEC; however, some of the phage types were identical. Determining whether certain Stx phages are associated with disease severity, however, requires further study.

In summary, the LPCR-RFLP method designed here for Stx phage genotyping can serve as a simple, rapid tool that provides high resolution and high reproducibility. We propose this strategy to better understand these mobile genetic elements that drive the evolution of STEC in the environment and reservoir species. By establishing a direct classification system, the frequency and dispersal of each Stx phage type can be monitored in specific STEC strain populations, thereby providing a useful tool for epidemiological studies. Additional future questions include better understanding the transduction and mutation frequency of Stx phages and how to prevent those phages from toxin production, which might help to approach the answer of decreasing STEC-associated diseases in human.

# APPENDIX

**Table 1**. STEC strain characteristics among 56 animals from four cattle herds and one deer population.

| Herd | Type of herd | STEC prevalence[b] | No. animals tested | No. strains[c] | No. unique strains[d] | Visit[e] | No. unique strains | Stx phage types | Serotypes |
|---|---|---|---|---|---|---|---|---|---|
| 8 | Beef | 50% | 22 | 112 | 50 | I | 20 | PT1, PT3, PT4 | O6:H34, O26, O157, NT |
| | | | | | | II | 11 | PT2, PT4 | O6:H34, O103:H2, NT |
| | | | | | | III | 12 | PT3, PT4 | O6: H34, O26:H11, O157:H7, NT |
| | | | | | | IV | 7 | PT3, PT4 | O26, NT |
| 9 | Dairy | 16% | 10 | 23 | 11 | I | 10 | PT5, PT6, PT7, PT8, PT10 | O157, NT |
| | | | | | | IV[f] | 1 | PT9 | NT |
| 11 | Beef | 7% | 6 | 11 | 7 | I | 4 | PT12, PT13 | O157, NT |
| | | | | | | II | 2 | PT12 | O157 |
| | | | | | | III | 1 | PT12 | O157 |
| 12 | Beef | 31% | 14 | 20 | 14 | I | 14 | PT5, PT10, PT14, PT15, PT16 | O157, NT |
| Deer[a] | wild | 11% | 4 | 8 | 4 | I | 4 | PT6, PT11 | O103, O157, NT |
| Total | | | 56 | 174 | 86 | | | 16 | |

**Table 1 (cont'd).**

[a] Deer fecal pellets were collected near the dairy pasture

[b] STEC prevalence in animals

[c] Number of STEC strains examined by phage typing and virulence gene profiling.

[d] Number of unique STEC strains characterized by different virulence profiles (*stx1*, *stx2*, eae) and O types. STEC isolated from different animals were regarded as unique strains.

[e] A subset of animals were followed at two week intervals for a total of four visits to identify the persistent shedders

[f] Stx phages were not recovered in visit II and III

**Table 2**. Primer sequences for serotype identification and *stx* profiling

| O type / gene | Primers | Oligonucleotide sequences (5'-3') | Amplicon (bp) | Reference |
|---|---|---|---|---|
| O26 | wzy26_F223 | TTAGGCGGTACCCATGAAGTCA | 242 | This study |
| | wzy26_R464 | GGTGCCATAAAGACAAAACAAAGA | | |
| O45 | wzy45_F188 | ATCGCGTTCGTCTGGATGAAAT | 443 | This study |
| | wzy45_R630 | AGCGCCCCTGATATCTCCTACAG | | |
| O103 | wzy103_F929 | CCCCGCGGGTATTTGCTAT | 184 | This study |
| | wzy103_R1112 | TCGTATGCGTTCGTTCTAAGATAA | | |
| O111 | wzy111_F495 | TTCCGTAATTTGCATCCTGATAC | 549 | This study |
| | wzy111_R1043 | TTTGCAAATCCATAAACAACTCC | | |
| O121 | wzy121_F306 | TACAGCCGGTAGTGTTGAAAGGAT | 626 | This study |
| | wzy121_R931 | CGCCCGTGTTAATATTCCAAGTC | | |
| O145 | wzy145_F754 | ATGGGCAGTATCTCTGGTATTGAA | 334 | This study |
| | wzy145_R1087 | TTGAAAGCCCGGATATTAGGAA | | |
| O157 | wzy157_F586 | GTAGGGGTTGTATGCTCGTTGTT | 278 | This study |
| | wzy157_R863 | GCTCCCATGTCTCCAAATACTTGT | | |
| *mdh* | mdh_F41 | AGGCGCTTGCACTACTGTTA | 835 | This study |
| | mdh_R875 | AGCGCGTTCTGTTCAAATG | | |
| *stx1* | Stx1-F | CGATGTTACGGTTTGTTACTGTGA | 242 | Muller D *et al*, 2007 |
| | Stx1-R | AATGCCACGCTTCCCAGAATTG | | |
| *stx2* | Stx2-F | GTTTTGACCATCTTCGTCTGATTATTG | 543 | Muller D *et al*, 2007 |
| | Stx2-R | AGCGTAAGGCTTCTGCTGTGAC | | |

**Table 2 (cont'd).**

| O type / gene | Primers | Oligonucleotide sequences (5'-3') | Amplicon (bp) | Reference |
|---|---|---|---|---|
| *Stx2a* | Stx2-1072F | AGGATGACACATTTACAGTGAAGGTT | 126 | Shringi S *et al*, 2012 |
| | Stx2-1197R | CACAGGTACTGGATTTGATTGTGAC | | |
| *stx2c* | Stx2c-F858 | CGACAGGCCCGTTATAAAAA | 243 | Shringi S *et al*, 2012 |
| | Stx2c-R1100 | GGCCACTTTTACTGTGAATGTATC | | |

**Table 3**. Primer sequences for Stx phage long range PCR

| Region | Gene | Primers | Oligonucleotide sequences (5'-3') | amplicon (bp) | Reference |
|---|---|---|---|---|---|
| Va | *q* | Q-F1 | GCATGGGCGGCAAATAAC | 1,759[a] | This study |
| | | Q-F2[c] | AATTCATGGAGAGCGTGGAG | | This study |
| Vb | *ninG* | ninG-F | ATCGGGCAGGAAGCAGTAGACG | 2,116[a] | This study |
| Vc | *roi* | Roi-F | GGAATTCACCGCTCGCCTTGTT | 3,055[a] | This study |
| | *stx1* | Stx1-R | AATGCCACGCTTCCCAGAATTG | | Muller D *et al*, 2007 |
| | *stx2* | Stx2-R | AGCGTAAGGCTTCTGCTGTGAC | | Muller D *et al*, 2007 |
| | *stx2c* | Stx2c-R1100 | GGCCACTTTTACTGTGAATGTATC | | Shringi S *et al*, 2012 |
| Vd | *ciii* | ciiiV-F1 | GACATTGCTCAGTGTATTCACTCGTTGGAA | | This study |
| | | ciiiV-F2[c] | GACATTGCTCCGTGTATTCACTCGTTGGAA | | Shima K *et al*, 2004 |
| | *stx1* | Stx1-longR | CAAACAAATTATCCCCTGTGCCACTA | 13,065[b] | This study |
| | *stx2a* | Stx2-longR | TATACGGACAGAGATATCGACCCCTC | 10,409[a] | This study |
| Via | *stx2a* | Bov stx2-F | GTGCCTGTTACTGGGTTTTTCTTC | | Paton AW, 1998 |
| | *j* | VTTF-R2 | TCTGTATCTGCCGGAAATCTGTCA | 16,805[a] | This study |

[a] The size of the amplicon is predicted based on the Stx2a phage in *E.coli* O157:H7 EDL933

[b] The size of the amplicon is predicted based on the Stx1 phage in *E.coli* O157:H7 EDL933

[c] Multiple primers were developed for *q* and *ciii* because of variation within the Stx phages. Q-F1 was used for Stx1 and Stx2c phages and Q-F2 was specifically used for the Stx2a phage; ciiiV-F1 was specifically used for the Stx1 phage and ciiiV-F2 was used for the Stx2a and Stx2c phages.

**Table 4**. Published full length Stx phage reference strains utilized in this study

| Type | Stx Phage | Accession number | Vd size (kb) | Via size (kb) |
|---|---|---|---|---|
| Stx1 | Prophage CP-933V | AE005174 | 10,408 | 29,530 |
| | Stx1 converting phage | NC004913 | 11,136 | 18,254 |
| | Enterobacteria phage VT1-Sakai | AP000400 | 10,408 | 28,497 |
| | Phage BP-4795 | AJ556162 | 12,278 | 29,616 |
| | Enterobacteria phage YYZ-2008 | FJ184280 | 11,975 | 30,638 |
| | Bacteriophage CP-1639 | AJ304858 | --[a] | 24,356 |
| Stx2a | Bacteriophage 933W | AF125520 | 13,065 | 16,805 |
| | Enterobacteria phage VT2-Sakai | AP000363 | 12,557 | 16,262 |
| | Enterobacteria phage VT2phi_272 | HQ424691 | --[a] | 18,098 |
| | Stx2 converting phage I | NC003525 | 13,065 | 16,811 |
| | Stx2 converting phage II | AP005154 | 12,557 | 18,119 |
| | Stx2-converting phage 86 | AB255436 | 13,309 | 17,474 |
| | Escherichia phage TL-2011c | JQ011316 | --[a] | 16,283 |
| | Escherichia phage P13374 | HE664024 | --[a] | 17,067 |
| | Enterobacteria phage Min 27 | EU311208 | 14,769 | 16,779 |
| | Stx2 converting phage vB_EcoP_24B | HM208303 | --[b] | 16,371 |
| Stx2c | Stx2-converting phage 1717 | NC011357 | 17,200 | 31,505 |
| | Enterobacteria phage 2851 | FM180578 | 15,848 | 30,222 |

[a] *In silico* analayis of the genome demonstrated missing *ciii* in some strains

[b] *In silico* analayis of the genome demonstrated an insertion mutation in *stx*

**Table 5**. Distribution of 16 distinct Stx phage types in four cattle herds and deer identified in present study

| Stx phage type | No. of strains that originated in | | | | | *stx* profile |
|---|---|---|---|---|---|---|
| | **8B** | **9D** | **11B** | **12B** | **Deer** | |
| **PT1** | 1 (2%) | | | | | ***stx1*** |
| **PT2** | 6 (12%) | | | | | ***stx1*** |
| **PT3** | 12 (24%) | | | | | ***stx1*** |
| **PT4** | 35 (70%) | | | | | ***stx2c*** |
| **PT5** | | 2 (18.2%) | | 10 (71.4%) | | ***stx1*** |
| **PT6** | | 6 (54.5%) | | | 2 (50%) | ***stx1*** |
| **PT7** | | 1 (9.1%) | | | | ***stx1*** |
| **PT8** | | 1 (9.1%) | | | | ***stx1*** |
| **PT9** | | 1 (9.1%) | | | | ***stx1*** |
| **PT10** | | 1 (9.1%) | | 4 (28.6%) | | ***stx2a*** |
| **PT11** | | | | | 2 (50%) | ***stx1*** |
| **PT12** | | | 6 (85.7%) | | | ***stx2c*** |
| **PT13** | | | 1 (14.3%) | | | ***stx2c*** |
| **PT14** | | | | 1 (7.1%) | | ***stx1*** |
| **PT15** | | | | 7 (50%) | | ***stx2a*** |
| **PT16** | | | | 2 (14.3%) | | ***stx2a*** |
| **Total** | 50 | 11 | 7 | 14 | 4 | 86 |

**Table 6**. Map of Stx phage types and serotypes for a subset of 41 unique STEC strains from four herds and four deer at visit I.

| Herd | Animal | *stx1* | *stx2a* | *stx2c* | *eae* | Serotype |
|---|---|---|---|---|---|---|
| **8B** | **1** | - | - | PT4 | - | NT[a] |
| | **2**[b] | - | - | PT4 | - | NT |
| | | f+[c] | - | PT4 | - | O6:H34[d] |
| | | f+[c] | - | PT4 | + | NT |
| | **3**[b] | - | - | PT4 | - | NT |
| | | f+[c] | - | PT4 | + | NT |
| | **5** | - | - | PT4 | - | NT |
| | **7** | - | - | PT4 | - | NT |
| | **8** | - | - | PT4 | - | NT |
| | **9**[b] | - | - | PT4 | - | NT |
| | | - | - | PT4 | + | NT |
| | **11** | - | - | PT4 | - | O6:H34[d] |
| | **13**[b] | - | - | PT4 | - | O6:H34[d] |
| | | - | - | PT4 | + | NT |
| | **14** | - | - | PT4 | + | NT |
| | **16**[b] | - | - | PT4 | - | NT |
| | | - | - | PT4 | + | O6:H34[d] |
| | **17** | - | - | PT4 | + | NT |
| | **19** | PT1[e] | - | - | - | NT |
| | **21** | PT3 | - | - | + | O26 |
| **9D** | **23** | PT5[f] | PT10[f] | - | + | O157 |
| | **25** | PT6 | - | - | + | NT |
| | **26** | PT7[e] | - | - | - | NT |
| | **27** | PT6 | - | - | + | NT |
| | **28** | PT6 | - | - | + | NT |
| | **29** | PT8[e] | - | - | - | NT |
| | **30**[b] | PT6 | - | - | - | NT |
| | | PT6 | - | - | + | NT |
| | **31** | PT5[d] | - | - | + | NT |
| | **32** | PT6 | - | - | + | NT |
| **Deer** | **33** | PT6 | - | - | + | NT |
| | **34** | PT6 | - | - | + | NT |
| | **35** | PT11 | - | - | + | O103 |
| | **36** | PT11 | - | - | + | O157 |

**Table 6 (cont'd)**.

| Herd | Animal | *stx1* | *stx2a* | *stx2c* | *eae* | Serotype |
|---|---|---|---|---|---|---|
| **11B** | **37** | - | - | PT12 | + | O157 |
| | **39** | - | - | PT13 | + | NT |
| | **40** | - | - | PT12 | + | O157 |
| | **41** | - | - | PT12 | + | O157 |
| **12B** | **43** | PT5 | PT15 | - | + | O157 |
| | **44** | - | PT15 | - | + | O157 |
| | **45** | PT5 | PT15 | - | + | O157 |
| | **46** | PT5 | PT10 | - | + | O157 |
| | **47** | PT5 | PT10 | - | + | O157 |
| | **48** | PT5 | PT15 | - | + | O157 |
| | **49** | PT5 | PT15 | - | + | O157 |
| | **50** | PT5 | PT10 | - | + | O157 |
| | **51** | PT5 | PT10 | - | + | O157 |
| | **52** | - | PT16 | - | + | NT |
| | **53** | PT5 | PT15 | - | - | O157 |
| | **54** | PT14 | - | - | + | NT |
| | **55** | PT5 | PT15 | - | + | O157 |
| | **56** | - | PT16 | - | + | NT |

[a] NT: Nontypeable using multiplex PCR assay targeting O157 and six non-O157 serotypes (O26, O45, O103, O111, O121, O145)

[b] Multiple STEC strains with unique virulence profiles were isolated from the same animal

[c] Faint positive band in *stx* multiplex PCR, but the Stx phage genome was not amplified in LPCR.

[d] Serotype results were determined using a SNP typing array and provided by the FDA for a subset of 16 strains.

[e] DNA extracted from regrowth culture of original single colony picks.

[f] P5 and P10 displayed the same RFLP patterns as Stx1 and Stx2a phages, respectively, in STEC O157:H7 Sakai outbreak strain.

**Table 7**. Map of Stx phage types and serotypes for a subset of unique STEC strains from 16 cattle in three herds at four sampling times (visits).

| | | Visits | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | I | | | II | | | III | | | IV | | |
| Herd | Animal | stx1 | stx2c | Serotype | stx1 | stx2c | Serotype | stx1 | stx2c | Serotype | stx1 | stx2c | Serotype |
| 8 | 4 | | | | | | | | | | PT3 | - | O26 |
| | 5[a] | - | PT4 | NT | PT2 | - | O103:H2[b] | - | PT4 | O6:H34[b] | - | PT4 | NT |
| | | | | | PT2 | PT4 | NT | | | | | | |
| | | | | | PT2 | PT4 | O103 | | | | | | |
| | 6[a] | | | | PT2 | - | O103:H2[b] | PT3 | - | O26:H11[b] | | | |
| | | | | | - | PT4 | NT | | | | | | |
| | 10 | | | | PT2 | PT4 | NT | PT3 | - | O26:H11[b] | | | |
| | 11[a] | - | PT4 | O6:H34[b] | | | | PT3 | - | O26:H11[b] | | | |
| | | | | | | | | PT3 | PT4 | O157:H7[b] | | | |
| | 12 | | | | - | PT4 | NT | | | | - | PT4 | NT |
| | 13[a] | - | PT4 | O6:H34[b] | | | | PT3 | - | O26 | | | |
| | | | | | | | | - | PT4 | O6:H34[b] | - | PT4 | NT |
| | 15[a] | | | | PT2 | - | O103:H2[b] | PT3 | - | O26:H11[b] | | | |
| | | | | | - | PT4 | O6:H34[b] | | | | | | |
| | 16[a] | - | PT4 | O6:H34[b] | - | PT4 | O6:H34[b] | PT3 | - | O26 | - | PT4 | NT |
| | | - | PT4 | NT | | | | PT3 | f+[c] | O26 | | | |
| | 18 | | | | | | | PT3 | - | O26 | | | |
| | 20[a] | | | | | | | - | PT4 | NT | PT3 | - | O26 |
| | | | | | | | | | | | - | PT4 | NT |
| | 22 | | | | - | PT4 | NT | | | | | | |

**Table 7 (cont'd)**.

| | | Visits | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | I | | | II | | | III | | | IV | | |
| **Herd** | **Animal** | **stx1** | **stx2c** | **Serotype** | **stx1** | **stx2c** | **Serotype** | **stx1** | **stx2c** | **Serotype** | **stx1** | **stx2c** | **Serotype** |
| **9** | **24** | | | | | | | | | | PT9 | | NT |
| **11** | **37** | - | PT12 | O157 | | | | - | PT12 | O157 | | | |
| | **38** | | | | | | | - | PT12 | O157 | | | |
| | **42** | | | | | | | | | | - | PT12 | O157 |

[a] Multiple STEC strains with unique virulence profiles were isolated from the same animal at one of the visits

[b] Serotype results were determined using a SNP typing array and provided by the FDA for a subset of 16strains.

[c] Faint positive band in *stx* multiplex PCR, but was not able to recover in Stx phage LPCR.

**Table 8**. Stx phage types in a set of 22 clinical STEC O157 strains recovered from patients with disease.

| Clade | Strain | Strain source | Year | stx1 | stx2a | stx2c |
|---|---|---|---|---|---|---|
| **1** | TW8264 | Sakai outbreak strain | 1996 | PT5 | PT10 | - |
| **2** | TW9084 | Michigan Patient | 2003 | PT5 | PT10 | - |
| | TW9088 | Michigan Patient | 2003 | PT5 | PT10 | - |
| | TW9089 | Michigan Patient | 2003 | PT5 | PT10 | - |
| | TW9124 | Michigan Patient | 2003 | PT5 | - | - |
| | TW9200 | Michigan Patient | 2003 | PT5 | Φ1 | - |
| | TW9375 | Michigan Patient | 2004 | - | PT10 | - |
| | TW9379 | Michigan Patient | 2004 | - | Φ2 | - |
| | TW14293 | Michigan Patient | 2005 | - | PT10 | - |
| | TW14297 | Michigan Patient | 2006 | PT5 | PT10 | |
| **7** | TW9087 | Michigan Patient | 2003 | Φ11 | - | Φ7 |
| | TW9178 | Michigan Patient | 2003 | - | - | Φ8 |
| | TW9195 | Michigan Patient | 2003 | Φ11 | - | Φ7 |
| | TW10119 | Michigan Patient | 2004 | - | - | Φ9 |
| | TW11067 | Michigan Patient | 2002 | - | - | PT12[a] |
| **8** | TW9085 | Michigan Patient | 2003 | - | Φ3 | Φ9 |
| | TW9378 | Michigan Patient | 2004 | - | Φ4 | Φ9 |
| | TW11050 | Michigan Patient | 2002 | - | Φ3 | Φ9 |
| | TW11070 | Michigan Patient | 2004 | - | P10 | Φ9 |
| | TW11507 | Michigan Patient | 2005 | P5 | Φ5 | - |
| | TW14307 | Michigan Patient | 2006 | - | Φ3 | Φ10 |
| | TW14359 | Spinach outbreak strain | 2006 | - | Φ6 | Φ9 |

[a] STEC O157 strain TW11067 carried the same Stx2c PT12 identified in beef herd 11.

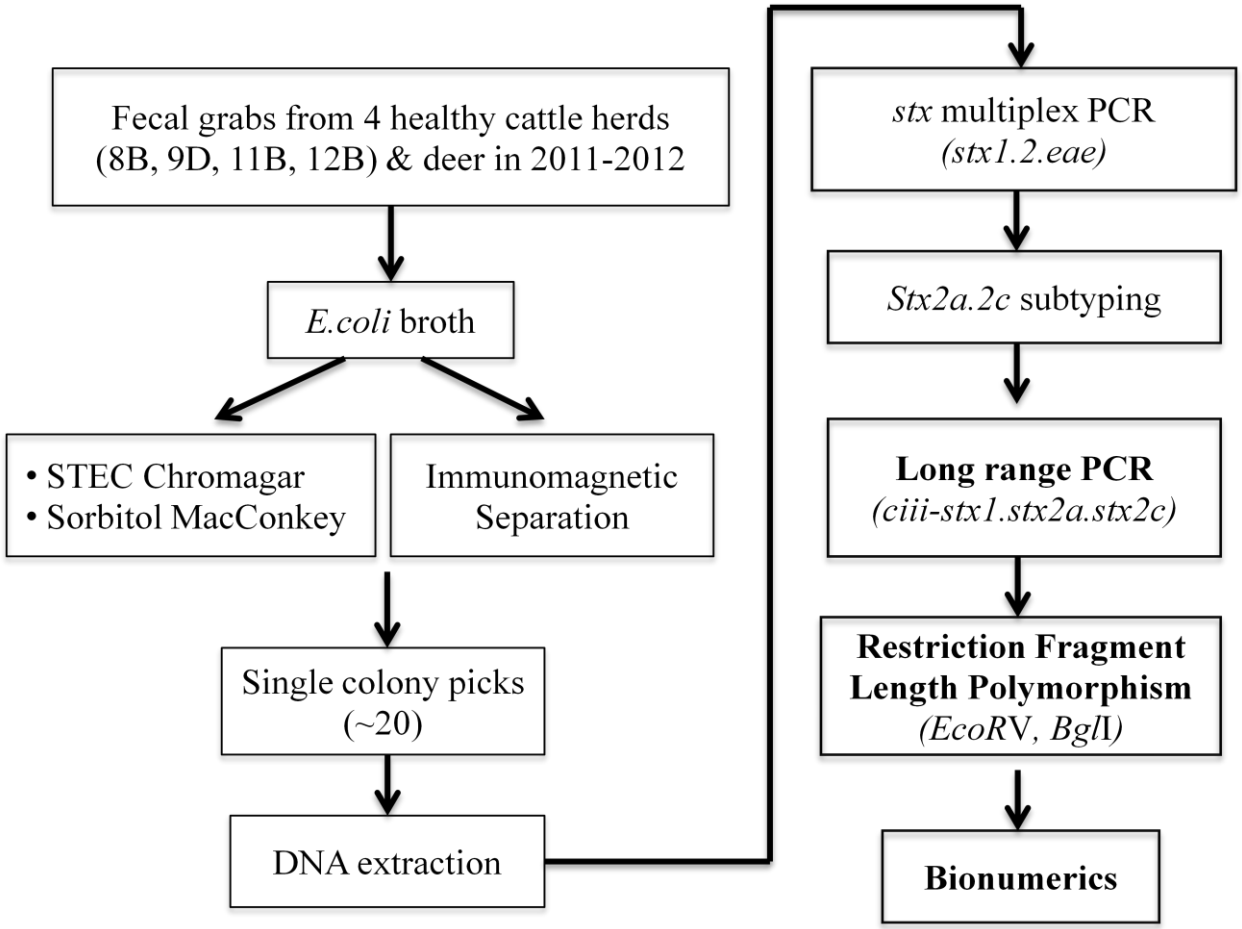


**Figure 1**. Flow chart illustrating how fecal samples from cattle and deer were processed for STEC isolation and Stx phage typing.

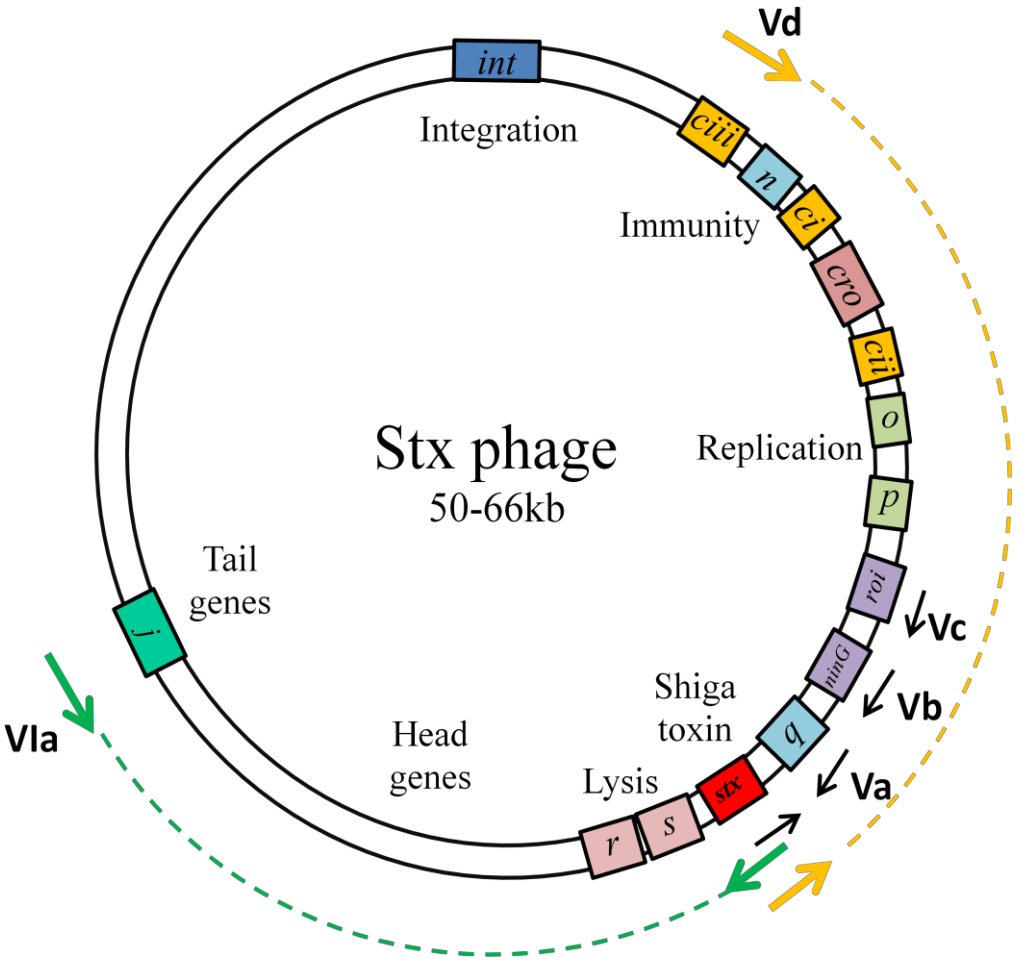


**Figure 2**. Regions targeted within the Stx phage genome for long range PCR (LPCR). Regions Va, Vb, Vc and Vd indicate four upstream regions of *stx* that extend to the Q, NinG, Roi, and CIII genes, respectively with amplicons of 1.6 kb, 2 kb, 3 kb and 12 kb in size. Region Via indicates the downstream region of *stx* that extends to the J gene encoding the tail fiber protein. The amplicon for region VIa is 16 to 30 kb.

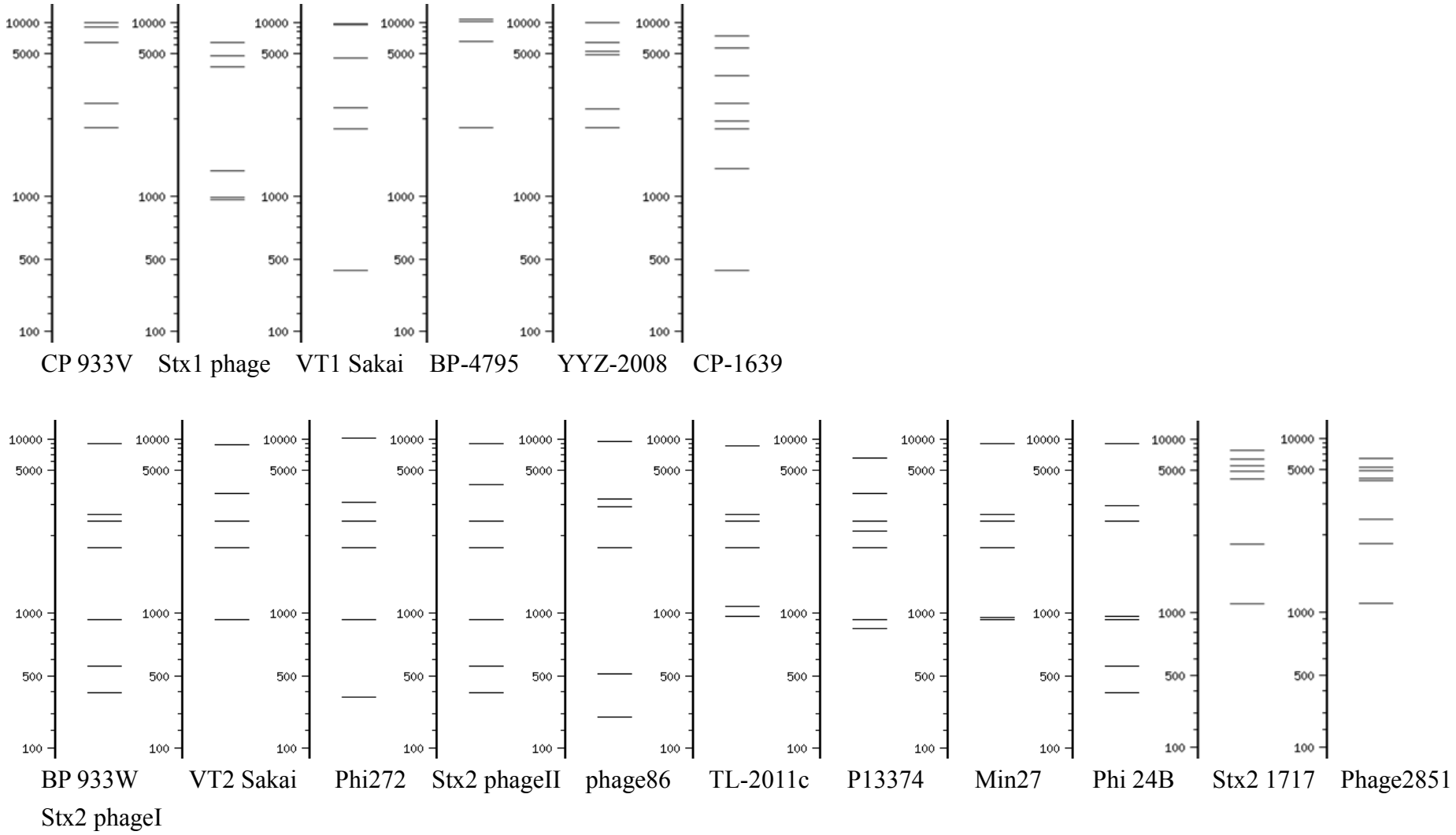


**Figure 3**. Stx phage *in silico* long range PCR based-RFLP patterns for region VIa of 18 reference strains with genome data available in NCBI. Restriction patterns for Stx2a phages are predicted following digestion with *Pst*I, while patterns for Stx1 and Stx2c phages are based on *Bcl*I digestion.

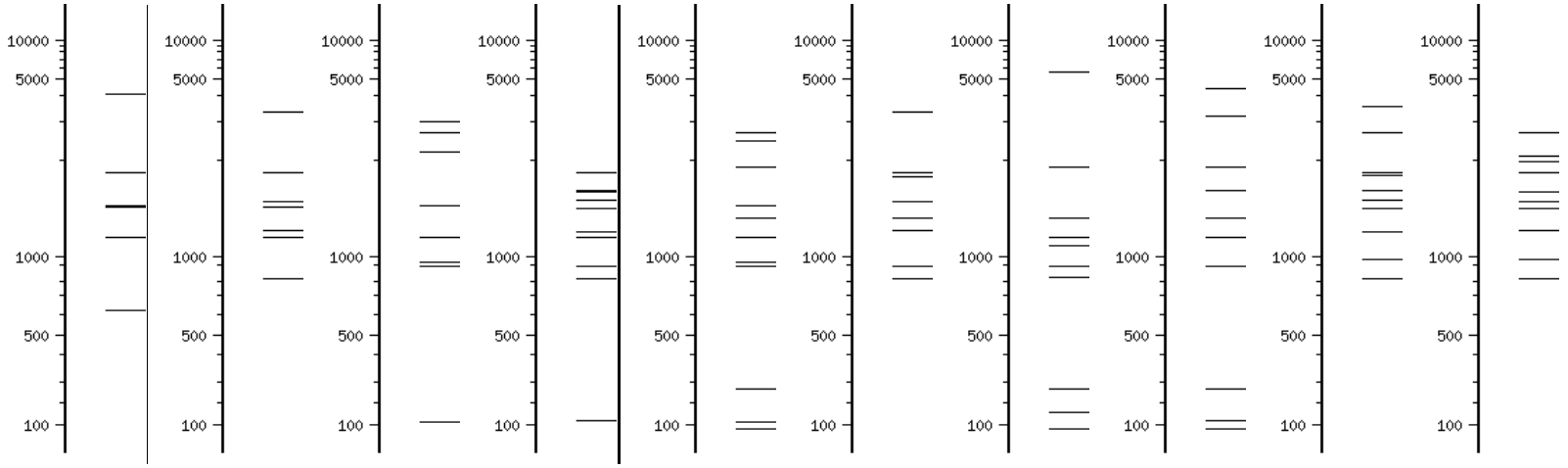


**Figure 4**. *In silico* prediction of Stx phage fragments in region Vd following long range PCR and digestion with *EcoR*V for 13 reference strains

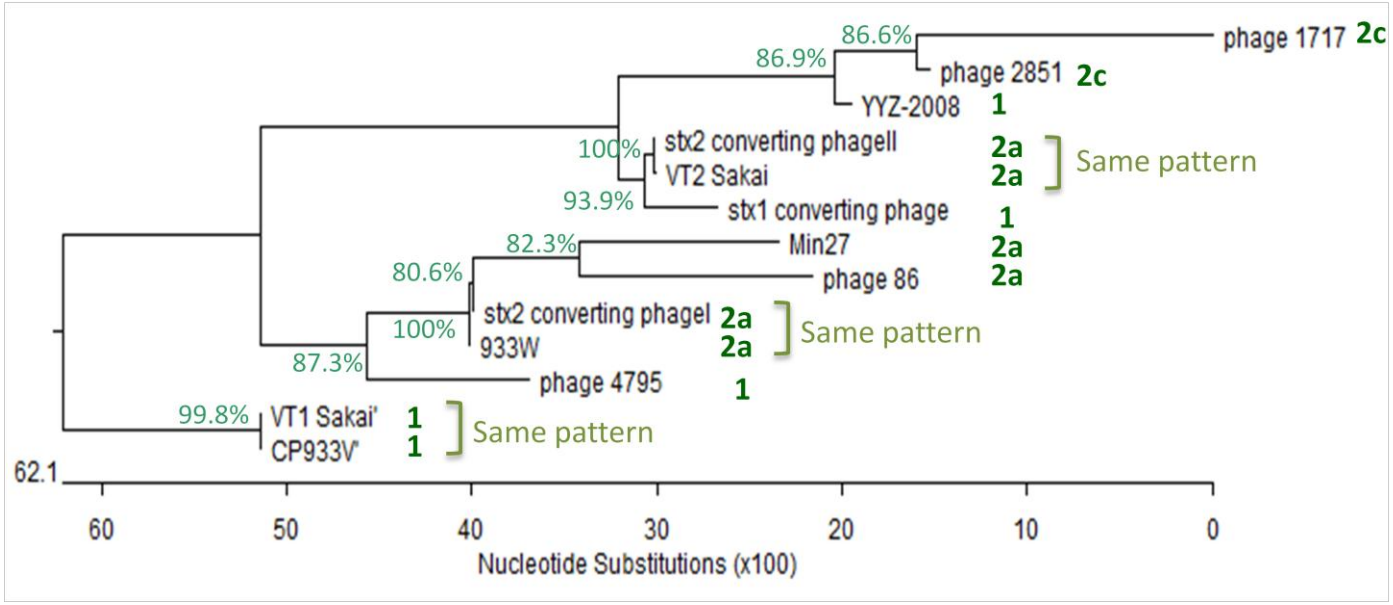


**Figure 5**. Phylogenetic analysis of Stx phage sequences specific for region Vd in 13 reference strains.

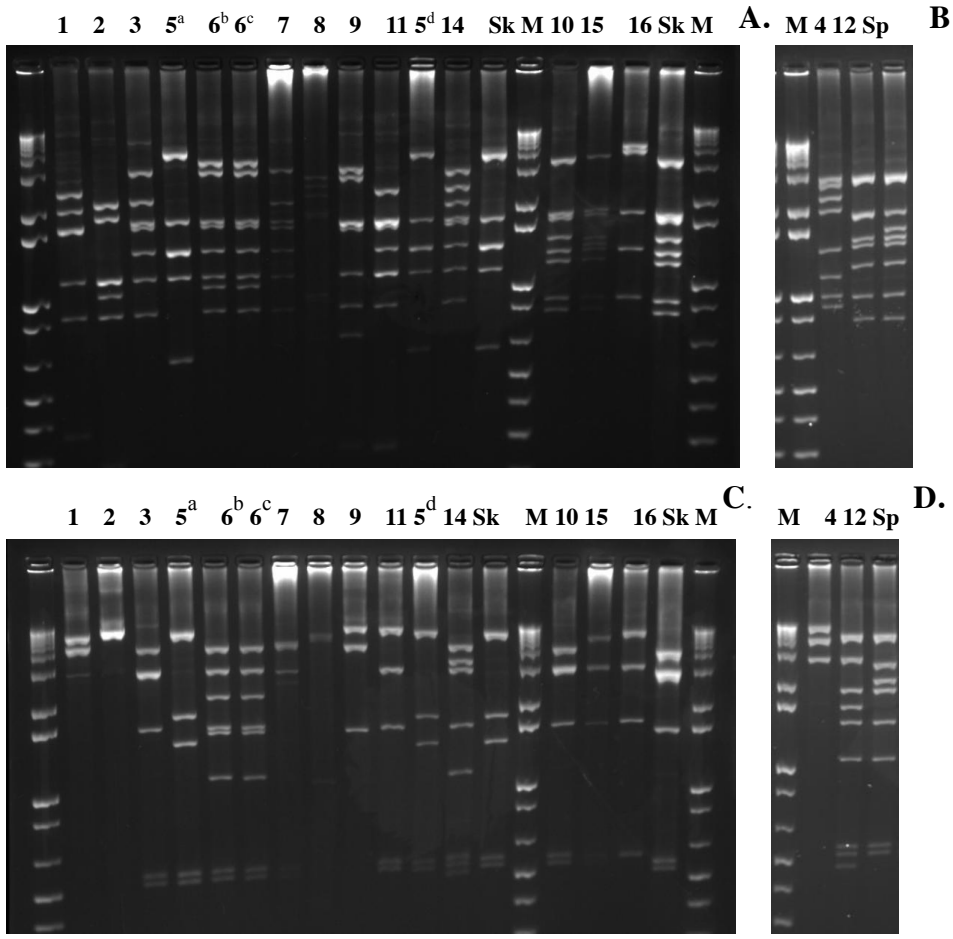


**Figure 6**. Long range PCR based-RFLP patterns in region Vd from 17 animals in four cattle herds and 1 deer. Panel A and B are showing restriction patterns based on *EcoR*V digestion; Panel C and D are showing restriction patterns based on *Bgl*I digestion. The Stx phage types (PTs) are listed on the top of each lane. Sk: STEC O157:H7 Sakai strain. Sp: STEC O157:H7 Spinich strain. In panel A and C, lane 2 to 13 belong to Stx1 phage and lane 15 to 18 belong to Stx2a phage. Panel B and D are showing Stx2c phages. [a]Stx1 phage PT5 from STEC strain in herd 9. [b]Stx1 phage PT6 from STEC strain in herd 9. [c]Stx1 phage PT6 from STEC strain in deer. [a]Stx1 phage PT5 from STEC strain in herd 12.

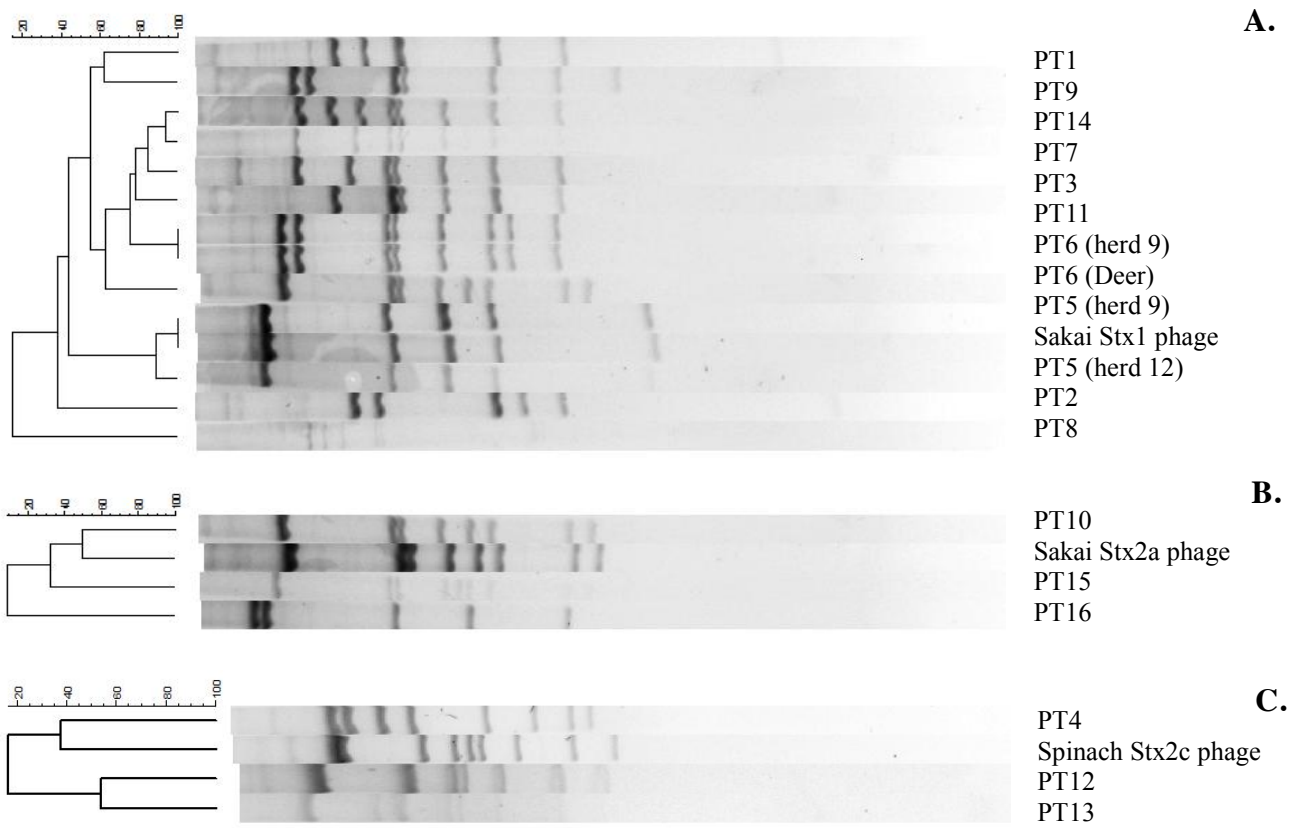


**Figure 7**. Dendrogram of Stx phage long range PCR based-RFLP patterns in region Vd generated by BioNumerics v 5.1 based on *EcoR*V digestion. The restriction banding patterns listed here represent all unique phage types identified in present study. Panel A is showing the restriction patterns of all 10 Stx1 phage types; panel B and C are showing all three Stx2a phage types and three Stx2c phage types, respectively.

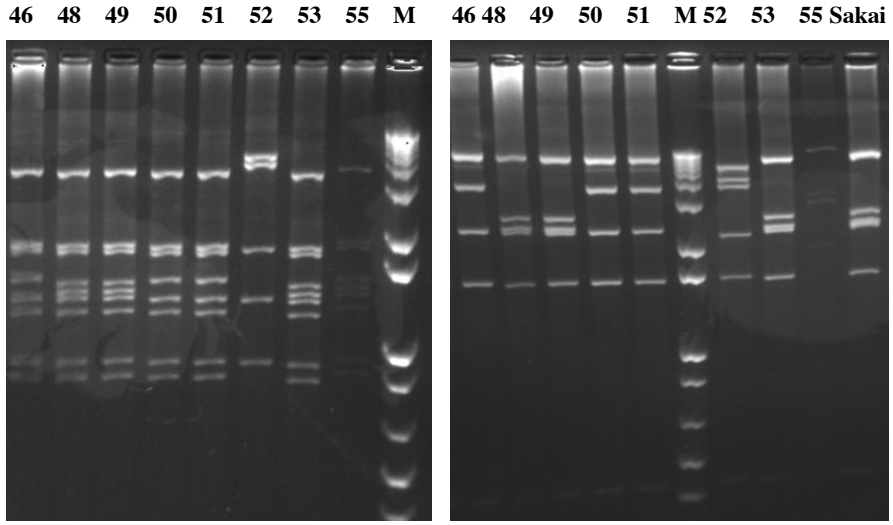


**Figure 8**. Comparison of Stx2a phage long range PCR based-RFLP results in region Vd (panel A) and VIa (panel B) for unique STEC strains in herd 12. The animal numbers from which each STEC strain was isolated are listed on the top of each lane. Panel A is showing restriction patterns in region Vd based on *EcoR*V digestion; Panel B is showing restriction patterns in region Via based on *Fsp*I digestion. The RFLP fragments from animal 55 was shifted upward due to the low DNA concentration of LPCR product. M: 1 kb plus marker

# BIBLIOGRAPHY

## BIBLIOGRAPHY


1. **Karmali MA.** 1989. Infection by verocytotoxin-producing *Escherichia coli*. Clin Microbiol Rev **2:**15-38.

2. **Cleary TG.** 2004. The role of shiga-toxin-producing *Escherichia coli* in hemorrhagic colitis and hemolytic uremic syndrome. Semin Pediatr Infect Dis **15:**260-265.

3. **Bielaszewska M, Schmidt H, Liesegang A, Prager R, Rabsch W, Tschäpe H, Cízek A, Janda J, Bláhová K, Karch H.** 2000. Cattle can be a reservoir of sorbitol-fermenting shiga toxin-producing *Escherichia coli* O157:H(-) strains and a source of human diseases. Journal of clinical microbiology **38:**3470-3473.

4. **Andreoli SP, Trachtman H, Acheson DW, Siegler RL, Obrig TG.** 2002. Hemolytic uremic syndrome: epidemiology, pathophysiology, and therapy. Pediatr Nephrol **17:**293-298.

5. **Ibarra C, Amaral MM, Palermo MS.** 2013. Advances in pathogenesis and therapy of hemolytic uremic syndrome caused by Shiga toxin-2. IUBMB life **65:**827-835.

6. **Endo Y, Tsurugi K, Yutsudo T, Takeda Y, Ogasawara T, Igarashi K.** 1988. Site of action of a Vero toxin (VT2) from *Escherichia coli* O157:H7 and of Shiga toxin on eukaryotic ribosomes. RNA N-glycosidase activity of the toxins. European journal of biochemistry / FEBS **171:**45-50.

7. **Trachtman H, Austin C, Lewinski M, Stahl RA.** 2012. Renal and neurological involvement in typical Shiga toxin-associated HUS. Nature reviews. Nephrology **8:**658-669.

8. **Fuller CA, Pellino CA, Flagler MJ, Strasser JE, Weiss AA.** 2011. Shiga toxin subtypes display dramatic differences in potency. Infect Immun **79:**1329-1337.

9. **Shringi S, Schmidt C, Katherine K, Brayton KA, Hancock DD, Besser TE.** 2012. Carriage of stx2a differentiates clinical and bovine-biased strains of *Escherichia coli* O157. PLoS One **7:**e51572.

10. **Friedrich AW, Bielaszewska M, Zhang WL, Pulz M, Kuczius T, Ammon A, Karch H.** 2002. *Escherichia coli* harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. The Journal of infectious diseases **185:**74-84.

11. **Louise CB, Obrig TG.** 1995. Specific interaction of *Escherichia coli* 0157:H7-derived Shiga-like toxin II with human renal endothelial cells. The Journal of infectious diseases **172:**1397-1401.

12. **Boerlin P, McEwen SA, Boerlin-Petzold F, Wilson JB, Johnson RP, Gyles CL.** 1999. Associations between virulence factors of Shiga toxin-producing *Escherichia coli* and disease in humans. Journal of clinical microbiology **37:**497-503.

13. **Beutin L, Krause G, Zimmermann S, Kaulfuss S, Gleier K.** 2004. Characterization of Shiga toxin-producing *Escherichia coli* strains isolated from human patients in Germany over a 3-year period. Journal of clinical microbiology **42:**1099-1108.

14. **Kawano K, Okada M, Haga T, Maeda K, Goto Y.** 2008. Relationship between pathogenicity for humans and stx genotype in Shiga toxin-producing *Escherichia coli* serotype O157. Eur J Clin Microbiol **27:**227-232.

15. **Nakamura H, Ogasawara J, Kita T, Hase A, Nishikawa Y.** 2008. Typing of Stx2 genes of *Escherichia coli* O157 isolates from cattle. Jpn J Infect Dis **61:**251-252.

16. **Steyert SR, Sahl JW, Fraser CM, Teel LD, Scheutz F, Rasko DA.** 2012. Comparative genomics and stx phage characterization of LEE-negative Shiga toxin-producing *Escherichia coli*. Frontiers in cellular and infection microbiology **2:**133.

17. **Beutin L, Hammerl JA, Strauch E, Reetz J, Dieckmann R, Kelner-Burgos Y, Martin A, Miko A, Strockbine NA, Lindstedt BA, Horn D, Monse H, Huettel B, Muller I, Stuber K, Reinhardt R.** 2012. Spread of a distinct Stx2-encoding phage prototype among *Escherichia coli* O104:H4 strains from outbreaks in Germany, Norway, and Georgia. Journal of virology **86:**10444-10455.

18. **Campbell A.** 1994. Comparative molecular biology of lambdoid phages. Annual review of microbiology **48:**193-222.

19. **Allison HE, Sergeant MJ, James CE, Jr S, Smith DL, Sharp RJ, Marks TS, McCarthy AJ.** 2003. Immunity profiles of wild-type and recombinant Shiga-Like toxin-encoding bacteriophages and characterization of novel double lysogens. Infect Immun **71:**3409-3418.

20. **Fogg PCM, Gossage SM, Smith DL, Saunders JR, McCarthy AJ, Allison HE.** 2007. Identification of multiple integration sites for Stx-phage Phi 24(B) in the

*Escherichia coli* genome, description of a novel integrase and evidence for a functional anti-repressor. Microbiol-Sgm **153:**4098-4110.

21. **Zhanga X, McDaniela AD, Wolfa LE, Keusch GT, Waldor MK, Acheson DWK.** 2000. Quinolone antibiotics induce shiga toxin-encoding bacteriophages, toxin production, and death in mice. 181 **2:**664-670.

22. **Wong CS, Jelacic S, Habeeb RL, Watkins SL, Tarr PI.** 2000. The risk of the hemolytic-uremic syndrome after antibiotic treatment of *Escherichia coli* O157:H7 infections. N Engl J Med **342:**1930-1936.

23. **McGannon CM, Fuller CA, Weiss AA.** 2010. Different classes of antibiotics differentially influence shiga toxin production. Antimicrob Agents Chemother **54:**3790-3798.

24. **Herold S, Siebert J, Huber A, Schmidt H.** 2005. Global expression of prophage genes in *Escherichia coli* O157:H7 strain EDL933 in response to norfloxacin. Antimicrob Agents Chemother **49:**931-944.

25. **Al-Jumaili I, Burke DA, Scotland SM, Al-Mardini H, Record CO.** 2006. A method of enhancing verocytotoxin production by *Escherichia coli*. FEMS Microbiol Lett **93:**121-125.

26. **Yue W-F, Du M, Zhu M-J.** 2012. High temperature in combination with UV irradiation enhances horizontal transfer of stx2 gene from *E. coli* O157:H7 to non-pathogenic *E. coli*. PLoS One **7:**1.

27. **Wagner PL, Livny J, Neely MN, Acheson DW, Friedman DI, Waldor MK.** 2002. Bacteriophage control of Shiga toxin 1 production and release by *Escherichia coli*. Molecular microbiology **44:**957-970.

28. **Wagner PL, Acheson DWK, Waldor MK.** 2001. Human neutrophils and their products induces higa toxin production by enterohemorrhagic *Escherichia coli*. Infect Immun **69:**1934-1937.

29. **Wagner PL, Neely MN, Zhang X, Acheson DW, Waldor MK, Friedman DI.** 2001. Role for a phage promoter in Shiga toxin 2 expression from a pathogenic *Escherichia coli* strain. J Bacteriol **183:**2081-2085.

30. **Neupane M, Abu-Ali GS, Mitra A, Lacher DW, Manning SD, Riordan JT.** 2011. Shiga toxin 2 overexpression in *Escherichia coli* O157:H7 strains associated with

severe human disease. Microbial pathogenesis **51:**466-470.

31. **Ritchie JM, Wagner PL, Acheson DWK, Waldor MK.** 2003. Comparison of Shiga toxin production by hemolytic-uremic syndrome-associated and bovine-associated Shiga toxin-producing *Escherichia coli* isolates. Appl Environ Microb **69:**1059-1066.

32. **Zhang YX, Laing C, Zhang ZZ, Hallewell J, You CP, Ziebell K, Johnson RP, Kropinski AM, Thomas JE, Karmali M, Gannon VPJ.** 2010. Lineage and Host Source Are Both Correlated with Levels of Shiga Toxin 2 Production by *Escherichia coli* O157:H7 Strains. Appl Environ Microb **76:**474-482.

33. **Asakura H, Makino SI, Kobori H, Watarai M, Shirahata T, Ikeda T, Takeshi K.** 2001. Phylogenetic diversity and similarity of active sites of Shiga toxin (Stx) in Shiga toxin-producing *Escherichia coli* (STEC) isolates from humans and animals. Epidemiology and infection **127:**27-36.

34. **Sullivan MB, Coleman ML, Quinlivan V, Rosenkrantz JE, DeFrancesco AS, Tan G, Fu R, Lee JA, Waterbury JB, Bielawski JP, Chisholm SW.** 2008. Portal protein diversity and phage ecology. Environ Microbiol **10:**2810-2823.

35. **Wang K, Chen F.** 2004. Genetic diversity and population dynamics of cyanophage communities in the Chesapeake Bay. Aquat Microb Ecol **34:**105-116.

36. **Smith DL, Wareing BM, Fogg PC, Riley LM, Spencer M, Cox MJ, Saunders JR, McCarthy AJ, Allison HE.** 2007. Multilocus characterization scheme for shiga toxin-encoding bacteriophages. Appl Environ Microbiol **73:**8032-8040.

37. **Johansen BK, Wasteson Y, Granum PE, Brynestad S.** 2001. Mosaic structure of Shiga-toxin-2-encoding phages isolated from *Escherichia coli* O157:H7 indicates frequent gene exchange between lambdoid phage genomes. Microbiology **147:**1929-1936.

38. **Volponi G, Rooks DJ, Smith DL, Picozzi C, Allison HE, Vigentini I, Foschino R, McCarthy AJ.** 2012. Short communication: Characterization of Shiga toxin 2-carrying bacteriophages induced from Shiga-toxigenic Escherichia coli isolated from Italian dairy products. Journal of dairy science **95:**6949-6956.

39. **Muniesa M, Blanco JE, De Simon M, Serra-Moreno R, Blanch AR, Jofre J.** 2004. Diversity of stx2 converting bacteriophages induced from Shiga-toxin-producing *Escherichia coli* strains isolated from cattle. Microbiology **150:**2959-2971.

40. **Khakhria R, Duck D, Lior H.** 1990. Extended Phage-Typing Scheme for Escherichia-Coli O157-H7. Epidemiology and infection **105:**511-520.

41. **Goh S, Newman C, Knowles M, Bolton FJ, Hollyoak V, Richards S, Daley P, Counter D, Smith HR, Keppie N.** 2002. *E. coli* O157 phage type 21/28 outbreak in North Cumbria associated with pasteurized milk. Epidemiology and infection **129:**451-457.

42. **Schmieger H.** 1999. Molecular survey of the Salmonella phage typing system of Anderson. J Bacteriol **181:**1630-1635.

43. **Smith DL, Rooks DJ, Fogg PCM, Darby AC, Thomson NR, McCarthy AJ, Allison HE.** 2012. Comparative genomics of Shiga toxin encoding bacteriophages. Bmc Genomics **13**.

44. **Kaiser AD, Jacob F.** 1957. Recombination between Related Temperate Bacteriophages and the Genetic Control of Immunity and Prophage Localization. Virology **4:**509-521.

45. **Echols H, Green L.** 1971. Establishment and Maintenance of Repression by Bacteriophage Lambda - Role of Ci, Cii, and Ciii Proteins. P Natl Acad Sci USA **68:**2190-&.

46. **Schubert RA, Dodd IB, Egan JB, Shearwin KE.** 2007. Cro's role in the CI-Cro bistable switch is critical for lambda's transition from lysogeny to lytic development. Gene Dev **21:**2461-2472.

47. **Sato T, Shimizu T, Watarai M, Kobayashi M, Kano S, Hamabata T, Takeda Y, Yamasaki S.** 2003. Distinctiveness of the genomic sequence of Shiga toxin 2-converting phage isolated from *Escherichia coli* O157:H7 Okayama strain as compared to other Shiga toxin 2-converting phages. Gene **309:**35-48.

48. **Herold S, Karch H, Schmidt H.** 2004. Shiga toxin-encoding bacteriophages--genomes in motion. International journal of medical microbiology : IJMM **294:**115-121.

49. **Hendrix RW, Lawrence JG, Hatfull GF, Casjens S.** 2000. The origins and ongoing evolution of viruses. Trends in microbiology **8:**504-508.

50. **Brussow H, Canchaya C, Hardt WD.** 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiology and

molecular biology reviews : MMBR **68:**560-602, table of contents.

51. **Smith DL, James CE, Sergeant MJ, Yaxian Y, Saunders JR, McCarthy AJ, Allison HE.** 2007. Short-tailed stx phages exploit the conserved YaeT protein to disseminate Shiga toxin genes among enterobacteria. Journal of bacteriology **189:**7223-7233.

52. **Rooks DJ, Yan Y, McDonald JE, Woodward MJ, McCarthy AJ, Allison HE.** 2010. Development and validation of a qPCR-based method for quantifying Shiga toxin-encoding and other lambdoid bacteriophages. Environmental microbiology **12:**1194-1204.

53. **Akiba M, Sameshima T, Nakazawa M.** 2000. Clonal turnover of enterohemorrhagic *Escherichia coli* O157:H7 in experimentally infected cattle. FEMS Microbiol Lett **184:**79-83.

54. **Mellmann A, Lu S, Karch H, Xu JG, Harmsen D, Schmidt MA, Bielaszewska M.** 2008. Recycling of Shiga toxin 2 genes in sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157 : NM. Appl Environ Microb **74:**67-72.

55. **Mellmann A, Bielaszewska M, Karch H.** 2009. Intrahost genome alterations in enterohemorrhagic *Escherichia coli*. Gastroenterology **136:**1925-1938.

56. **Imamovic L, Balleste E, Jofre J, Muniesa M.** 2010. Quantification of Shiga Toxin-Converting Bacteriophages in Wastewater and in Fecal Samples by Real-Time Quantitative PCR. Appl Environ Microb **76:**5693-5701.

57. **Smith-Vaughan HC, Sriprakash KS, Mathews JD, Kemp DJ.** 1995. Long PCR-ribotyping of nontypeable Haemophilus influenzae. Journal of clinical microbiology **33:**1192-1195.

58. **Shima K, Terajima J, Sato T, Nishimura K, Tamura K, Watanabe H, Takeda Y, Yamasaki S.** 2004. Development of a PCR-restriction fragment length polymorphism assay for the epidemiological analysis of Shiga toxin-producing *Escherichia coli*. Journal of clinical microbiology **42:**5205-5213.

59. **Shima K, Yoshii N, Akiba M, Nishimura K, Nakazawa M, Yamasaki S.** 2006. Comparison of PCR-RFLP and PFGE for determining the clonality of enterohemorrhagic Escherichia coli strains. FEMS microbiology letters **257:**124-131.

60. **Zeng YH, Koblizek M, Li YX, Liu YP, Feng FY, Ji JD, Jian JC, Wu ZH.** 2013.

Long PCR-RFLP of 16S-ITS-23S rRNA genes: a high-resolution molecular tool for bacterial genotyping. Journal of applied microbiology **114:**433-447.

61. **Manning SD, Motiwala AS, Springman AC, Qi W, Lacher DW, Ouellette LM, Mlaclonicky JM, Somsel P, Rudrik JT, Dietrich SE, Zhang W, Swaminathan B, Alland D, Whittam TS.** 2008. Variation in virulence among clades of *Escherichia coli* O157 : H7, associated with disease outbreaks. P Natl Acad Sci USA **105:**4868-4873.

62. **Dwivedi B, Schmieder R, Goldsmith DB, Edwards RA, Breitbart M.** 2012. PhiSiGns: an online tool to identify signature genes in phages and design PCR primers for examining phage diversity. BMC bioinformatics **13:**37.

63. **Dombek PE, Johnson LK, Zimmerley ST, Sadowsky MJ.** 2000. Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. Appl Environ Microb **66:**2572-2577.

64. **Unkmeir A, Schmidt H.** 2000. Structural analysis of phage-borne stx genes and their flanking sequences in shiga toxin-producing *Escherichia coli* and Shigella dysenteriae type 1 strains. Infect Immun **68:**4856-4864.

65. **Miyamoto H, Nakai W, Yajima N, Fujibayashi A, Higuchi T, Sato K, Matsushiro A.** 1999. Sequence analysis of Stx2-converting phage VT2-Sa shows a great divergence in early regulation and replication regions. DNA research : an international journal for rapid publication of reports on genes and genomes **6:**235-240.

66. **L'Abee-Lund TM, Jorgensen HJ, O'Sullivan K, Bohlin J, Ligard G, Granum PE, Lindback T.** 2012. The highly virulent 2006 Norwegian EHEC O103:H25 outbreak strain is related to the 2011 German O104:H4 outbreak strain. PLoS One **7:**e31413.

67. **Creuzburg K, Kohler B, Hempel H, Schreier P, Jacobs E, Schmidt H.** 2005. Genetic structure and chromosomal integration site of the cryptic prophage CP-1639 encoding Shiga toxin 1. Microbiology **151:**941-950.

68. **Gamage SD, Patton AK, Hanson JF, Weiss AA.** 2004. Diversity and host range of Shiga toxin-encoding phage. Infect Immun **72:**7131-7139.

69. **Wagner PL, Acheson DWK, Waldor MK.** 1999. Isogenic lysogens of diverse Shiga toxin 2-encoding bacteriophages produce markedly different amounts of Shiga toxin. Infect Immun **67:**6710-6714.

70. **Kim J, Nietfeldt J, Benson AK.** 1999. Octamer-based genome scanning distinguishes a unique subpopulation of *Escherichia coli* O157 : H7 strains in cattle. P Natl Acad Sci USA **96:**13288-13293.

71. **Park D, Stanton E, Ciezki K, Parrell D, Bozile M, Pike D, Forst SA, Jeong KC, Ivanek R, Dopfer D, Kaspar CW.** 2013. Evolution of the Stx2-Encoding Prophage in Persistent Bovine *Escherichia coli* O157:H7 Strains. Appl Environ Microb **79:**1563-1572.

72. **Chase-Topping ME, McKendrick IJ, Pearce MC, MacDonald P, Matthews L, Halliday J, Allison L, Fenlon D, Low JC, Gunn G, Woolhouse ME.** 2007. Risk factors for the presence of high-level shedders of *Escherichia coli* O157 on Scottish farms. Journal of clinical microbiology **45:**1594-1603.