Phylogeny of Shiga Toxin-Producing *Escherichia coli* O157 Isolated from Cattle and Clinically Ill Humans

James L. Bono,[1] Timothy P. L. Smith,[1] James E. Keen,[2] Gregory P. Harhay,[1] Tara G. McDaneld,[1] Robert E. Mandrell,[3] Woo Kyung Jung,[4] Thomas E. Besser,[4] Peter Gerner-Smidt,[5] Martina Bielaszewska,[6] Helge Karch,[6] and Michael L. Clawson*,[1]

[1]United States Department of Agriculture, Agricultural Research Service, US Meat Animal Research Center, Clay Center, Nebraska

[2]Great Plains Veterinary Educational Center, University of Nebraska

[3]United States Department of Agriculture, Agricultural Research Service, Western Regional Research Center, Albany, California

[4]Department of Veterinary Microbiology and Pathology, Washington State University

[5]Enteric Diseases Laboratory Response Branch, Centers for Disease Control and Prevention, Atlanta, Georgia

[6]Institute of Hygiene and the National Consulting Laboratory on Haemolytic Uraemic Syndrome, University of Münster, Münster, Germany

***Corresponding author:** E-mail: mike.clawson@ars.usda.gov.

Associate editor: Helen Piontkivska

Research article

Abstract

Cattle are a major reservoir for Shiga toxin-producing *Escherichia coli* O157 (STEC O157) and harbor multiple genetic subtypes that do not all associate with human disease. STEC O157 evolved from an *E. coli* O55:H7 progenitor; however, a lack of genome sequence has hindered investigations on the divergence of human- and/or cattle-associated subtypes. Our goals were to 1) identify nucleotide polymorphisms for STEC O157 genetic subtype detection, 2) determine the phylogeny of STEC O157 genetic subtypes using polymorphism-derived genotypes and a phage insertion typing system, and 3) compare polymorphism-derived genotypes identified in this study with pulsed field gel electrophoresis (PFGE), the current gold standard for evaluating STEC O157 diversity. Using 762 nucleotide polymorphisms that were originally identified through whole-genome sequencing of 189 STEC O157 human- and cattle-isolated strains, we genotyped a collection of 426 STEC O157 strains. Concatenated polymorphism alleles defined 175 genotypes that were tagged by a minimal set of 138 polymorphisms. Eight major lineages of STEC O157 were identified, of which cattle are a reservoir for seven. Two lineages regularly harbored by cattle accounted for the majority of human disease in this study, whereas another was rarely represented in humans and may have evolved toward reduced human virulence. Notably, cattle are not a known reservoir for *E. coli* O55:H7 or STEC O157:H$^-$ (the first lineage to diverge within the STEC O157 serogroup), which both cause human disease. This result calls into question how cattle may have originally acquired STEC O157. The polymorphism-derived genotypes identified in this study did not surpass PFGE diversity assessed by *Bln*I and *Xba*I digestions in a subset of 93 strains. However, our results show that they are highly effective in assessing the evolutionary relatedness of epidemiologically unrelated STEC O157 genetic subtypes, including those associated with the cattle reservoir and human disease.

Key words: STEC O157, phylogeny, evolution, disease, cattle, human.

Introduction

Shiga toxin-producing *Escherichia coli* O157 (STEC O157) are genetically diverse bacteria that cause diarrhea, hemorrhagic colitis, and hemolytic uremic syndrome (HUS) in humans (Griffin and Tauxe 1991; Gunzer et al. 1992). STEC O157:H7 and STEC O157:H$^-$ strains comprise the STEC O157 serogroup. STEC O157:H7 strains are typically motile, do not ferment sorbitol (SOR$^-$), and most do not express β-glucuronidase activity (GUD$^-$), whereas STEC O157:H$^-$ strains are typically nonmotile, SOR$^+$, and GUD$^+$ (fig. 1). Both subgroups are recently emerged pathogens; however, STEC O157:H$^-$ human infections have been primarily limited to Europe (Alpers et al. 2009; Pollock et al. 2010). In contrast, STEC O157:H7 is responsible for a large majority of human STEC O157 infections throughout the world (Griffin and Tauxe 1991; Mead and Griffin 1998). Thus, the STEC O157:H7 subgroup, in particular, is an international concern.

Cattle are a major reservoir for STEC O157:H7 and a source of human infection (Borczyk et al. 1987; Griffin and Tauxe 1991; Wells et al. 1991). However, cattle also harbor certain STEC O157:H7 genetic subtypes that are rarely found in clinically ill humans (Kim et al. 1999; Roldgaard et al. 2004; Besser et al. 2007; Bono et al. 2007; Clawson et al. 2009; Whitworth et al. 2010). Cattle occasionally harbor SOR$^+$ STEC O157:H$^-$, although the primary reservoir for this serotype is unknown (Bielaszewska et al. 2000; Lee and Choi 2006; Alpers et al. 2009). Consequently, cattle and human clinical STEC O157 cases represent important focal points for identifying STEC O157 genetic subtypes, understanding STEC O157 evolution and the evolution of genetic