Journal of Infection and Public Health 12 (2019) 884–889

Contents lists available at ScienceDirect

Journal of Infection and Public Health

journal homepage: http://www.elsevier.com/locate/jiph

Whole genome sequencing to characterize shiga toxin-producing *Escherichia coli* O26 in a public health setting

Baha Abdalhamid[a,*], Emily L. Mccutchen[b], Alyssa C. Bouska[a], Zhang Weiwei[a], Brianna Loeck[c], Steven H. Hinrichs[a], Peter C. Iwen[a]

[a] Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE, United States
[b] Nebraska Public Health Laboratory, University of Nebraska Medical Center, Omaha, NE, United States
[c] Division of Public Health – Epidemiology and Informatics Unit, Nebraska Department of Health and Human Services, Omaha, NE, United States

ARTICLE INFO

Article history:
Received 14 January 2019
Received in revised form 27 May 2019
Accepted 9 June 2019

Keywords:
Whole genome sequencing
Shiga toxin
E. coli O26
Public health

ABSTRACT

Background: Shiga-toxin producing *Escherichia coli* (STEC) O26:H11 is the second most common cause of severe diarrhea and hemolytic uremic syndrome worldwide. The implementation of whole genome sequencing (WGS) enhances the detection and in-depth characterization of these non-O157 STEC strains. The aim of this study was to compare WGS to phenotypic serotyping and pulse field gel electrophoresis (PFGE) for characterization of STECO26 strains following a zoonotic outbreak from cattle to humans.
Methods and results: This study evaluated seven *E. coli* strains; two strains isolated from two children with gastrointestinal symptoms and five strains from five calves suspected as the source of infection. Six of these isolates were serotyped phenotypically and by WGS as *E. coli* O26:H11 while one bovine isolate could be serotyped only by WGS as *E. coli* O182:H25. Stx1 was detected in two human- and two bovine-isolates using PCR and WGS. Using WGS, all four STECO26 isolates belong to sequence type (ST) 21 while the two stx1 negative *E. coli* O26 were ST29. All four STECO26 isolates were indistinguishable by PFGE. However, the data generated by WGS linked the two human STECO26 isolates to only one bovine STECO26 strain by having identical high-quality single nucleotide polymorphisms (hqSNPs) and identical virulence factor profiles while the remaining bovine STECO26 isolate differed by 7 hqSNPs and lacked virulence factor toxB.
Conclusions: These data demonstrated that WGS provided significant information beyond traditional epidemiological tools allowing for comprehensive characterization of the STEC. Using this approach, WGS was able to identify the specific source of infection in this study.

Introduction

Shiga toxin-producing *Escherichia coli* (STEC) causes a variety of gastrointestinal symptoms ranging from mild to severe diarrhea with or without blood [1,2]. STEC is one of seven foodborne pathogens evaluated by members of PulseNet, a national laboratory network that connects foodborne illnesses to detect outbreaks [3–6]. STEC, both those identified as serotype O157 and the non-O157 serotypes, is the 3rd or 4th most common foodborne pathogen detected [4–7]. According to the Centers for Diseases Control and Prevention (CDC), 6 non-O157 O groups, O26, O45, O103, O111, O121, and O145, are associated with ≥70% of non-O157 STEC diseases [6,8,9]. All STEC encode either one or both of the stx1 and stx2 genes leading to the production of the Stx1 or Stx2 shigatoxins that may cause in severe cases of disease, hemorrhagic colitis (HC) and/or hemolytic uremic syndrome (HUS) [1,10,11].

Although *E. coli* O157:H7 is the most common serotype causing HC and HUS, *E. coli* O26:H11 (STEC O26) is recognized as the most common non-O157 STEC serotype causing these disease syndromes [9,12,13]. Even though the production of Stx1 and/or Stx2 is a major virulence factor for STEC, other virulence factors are also known to be responsible for disease symptoms [11,14,15]. For instance, STEC O26 can encode for virulence factors that are significant causes of diseases such as the eae gene, ehxA gene, and espP gene which are responsible for attaching and effacing, hemolysin, and protease activity, respectively [12,14,15]. Based on these virulence factors, STEC O26 has been divided into two main groups, the Enteropathogenic *E. coli* (EPEC) which encodes the eae gene and the STEC which encodes stx1and/or stx2 genes in addition to

* Corresponding author at: DRC II 7035, Department of Pathology and Microbiology, University of Nebraska Medical Center, 985900 Nebraska Medical Center, Omaha, NE 68198-5900, United States.
E-mail address: babdalhamid@unmc.edu (B. Abdalhamid).

https://doi.org/10.1016/j.jiph.2019.06.008

the eae gene [10,12,16]. Moreover, STEC O26 can be separated into sequencing types (ST) using multilocus sequence typing (MLST), with ST21 and ST29 being the most common sequencing types associated with human diseases [10,12,17].

More recently, the CDC in consultation with federal, local, and state public health laboratories has been undergoing a shift in laboratory technology for outbreak investigating of foodborne diseases. Since 1998, pulse field gel electrophoresis (PFGE) has been the gold standard method used to evaluate isolates for strain relatedness. Although technically demanding, this method has been shown to be a reliable method [6,15,18]. Serotyping has also undergone changes from the standards slide agglutination test to utilization of Luminex-based technology [19]. Both methods require extensive training and expensive reagents, with PFGE recognized as having a limited ability to discriminate among outbreaks in some cases [15,18–20]. Most recently, whole genome sequencing (WGS) is becoming the gold standard for outbreak investigation with the ability not only to serotype STEC strains but to discern relatedness among isolates [12,15,20,21]. With the reduced cost and time to complete high throughput sequencing, WGS is rapidly becoming a standard use in epidemiological diagnostics for outbreak investigation and surveillance studies [15,18,20,21].

The aim of this study was to evaluate WGS technology as a routine public health laboratory method for the epidemiological investigation of an unusual outbreak of diarrhea caused by STEC O26. To achieve this aim, a comparative analysis among WGS, phenotypic serotyping, PCR, and PFGE was done.

Materials and methods

Outbreak investigation

The Nebraska State Department of Health and Human Services (DHHS) investigated two cases of STEC in 11-year-old female and 8-year-old male siblings. Both developed abdominal pain lasting for 5 days while the male also experienced bloody diarrhea. A multiplex PCR panel (FilmArray Gastrointestinal [GI] panel, Biofire Diagnostics, UT, USA) performed on stool specimens from each patient, detected a shiga-like toxin-producing *E. coli* (STEC) (non O157:H7) which subsequently resulted in the stool being forwarded to the Nebraska Public Health Laboratory (NPHL) for culture and further analysis. Local health officials interviewed the patients to obtain demographic, clinical, and exposure information and the results showed that each patient had contact with 4H calves which were also symptomatic with diarrhea. Five stool specimens from the associated 4H calves were subsequently requested by DHHS for testing. These specimens were cultured at the State Nebraska Veterinary Diagnostics Laboratory, Lincoln, NE with the isolated bovine *E. coli* strains of clinical significance sent to NPHL for epidemiological testing.

Bacterial strains

Seven *E. coli* isolates were collected overall, which included two human isolates, (one from each patient, coded # SRR6293238 and #SRR6293256) and five cattle isolates (coded # SRR58776420, # SRR6057057, # SRR6057923, # SRR5876423, and # SRR6057924). These isolates were submitted to the NPHL for additional testing, including serotyping using the slide agglutination test, PFGE, and WGS.

Slide agglutination test

The bovine- and human-associated *E. coli* strains were serotyped by means of a slide agglutination test utilizing antiserum specific for the following *E. coli* serotypes: O26, O45, O103, O111, O121, and O145 (Statens Serum Institut, Copenhagen, Denmark) as recommended by the manufacturer.

DNA extraction

Bacterial genomic DNA from all seven isolates was extracted from overnight cultures using the MagNA Pure Compact Nucleic Acid Isolation Kit I using the MagnaPure Compact Instrument (Roche Diagnostics, IN, USA). Qualitative and quantitative DNA measures were determined via the NanoDrop 2000 UV–vis spectrophotometer (Fisher, MA, USA) and Qubit® 3.0 Fluorometer (Invitrogen, CA, USA), respectively.

Pulse field gel electrophoresis

To determine *E. coli* strain relatedness, pulse field gel electrophoresis (PFGE) was performed based upon the Centers for Disease Control and Prevention PulseNet protocols. Briefly, extracted DNA was embedded in 1% agarose, lysed, and subjected to digestion with XbaI restriction enzyme [22]. The digested product was then separated on a 1% pulse field certified agarose gel using a CHEF Electrophoresis Cell (BioRad, CA, USA). The resulting pattern, or fingerprint, was analyzed with BioNumerics software v 6.6 (Applied Maths, Sint-Martens-Laten, Belgium) and compared to local and national databases of known *E. coli* PFGE fingerprints.

Whole genome sequencing

Whole genome sequencing (WGS) using a 250-bp paired end chemistry was performed on the bacterial genomic DNA on an Illumina MiSeq (Illumina, CA, USA) according to the manufacturer's instructions. Libraries were constructed using the Nextera XT Library Prep Kit (Illumina). Quality assessment of the run was determined using the following parameters: Phred quality score (QS30 > 75%), cluster density (600–1300), and clusters passing filters (>80%) [15]. After subsequent quality trimming, the contigs of genomic sequences were de novo assembled with a minimum contig size threshold of 200 bps using CLC Genomics Workbench v11.0.1 (Qiagen, CA, USA). The draft genome of each strain was annotated using Prokka Web-based tool v1.12 (https://github.com/tseemann/prokka).

Bioinformatics following whole genome sequencing

The serotype of each strain was determined using serotyping genes to include O-type (wzx and wzy) and H-type (fliC) deposited in the Center for Genomic Epidemiology (CGE) (http://www.genomicepidemiology.org). Serotype Finder 2.0 Web-based tool (https://cge.cbs.dtu.dk/services/SerotypeFinder/) identified the serotypes for the *E. coli* strains in FASTA format using the following parameters: 85% threshold for %ID and 60% minimum length (the number of nucleotides in a sequence of interest that must overlap a serotype gene to count as a hit for that gene) [20].

Virulence Finder 2.0 Web-based tool (https://cge.cbs.dtu.dk/services/VirulenceFinder/) was used to identify virulence genes in strains analyzed in this study with the following parameters: 90% threshold for %ID and 60% minimum length. Virulence genes for *E. coli* that had been deposited in CGE were used for this process [15].

MLST 2.0 Web-based tool (https://cge.cbs.dtu.dk/services/MLST/) was utilized for MLST analysis using 7 housekeeping *E. coli* genes (adk, fumC, gyrB, icd, mdh, purA, and recA) as previously described [17].

Mapping and high quality single nucleotide polymorphisms (hqSNPs) analysis were done using the lyve-SET (v.2.0.1) pipeline (https://github.com/lskatz/lyve-SET) [23]. Briefly, reads were

cleaned with the CGP read cleaner and sequences aligned to the *E. coli* O26:H11 strain 11368 genome (AP010953) with BWA (v0.7.17). Variants were called with Varscan (v2.3.7) and the Lyve-SET *E. coli* settings and mask-cliffs options were used. The resulting hqSNPs were then used for creating a SNP matrix. The UPGMA (Unweighted Pair Group Method with Arithmetic Mean) method was used to construct a phylogenetic tree and the output dendrogram was presented in Newick format. The Manhattan coefficient method was used to compare between sets of variables.

Detection of stx1/2 genes by PCR

To confirm WGS and multiplex PCR results, both stx genes were amplified by a traditional PCR method. Previously described stx gene primers specific for the amplification of both the stx1 and stx2 gene sequences were used [24].

Nucleotide sequence accession numbers

The genome sequences of all seven *E. coli* strains analyzed in this study were deposited in the GeneBank with the following accession numbers: SRR6057923, SRR5876423, SRR6293238, SRR6293256, SRR6057924, SRR5876420, and SRR6057057.

Results

Slide agglutination test

Six *E. coli* strains agglutinated with the antiserum specific for *E. coli* O26. One bovine strain (SRR6057957) did not agglutinate with any of the available antiserum: O26, O45, O103, O111, O121, and O145.

PFGE comparison and strain phylogenetic relatedness

Using BioNumerics software v 6.6, local and national known fingerprint comparisons indicated that six strains were *E. coli* serotype O26. The seventh strain, SRR6057057, did not match any known fingerprint locally or nationally by ≥90%. Comparison within the strains showed that the human strains (SRR6293238 and SRR6293256) as well as bovine strains (SRR6057923 and SRR5876423) had indistinguishable PFGE fingerprints. Bovine strains SRR6057924 and SRR5876420 were also indistinguishable from each other, but only matched at 82.9% to the four identical human and bovine strains. The fingerprint of bovine strain SRR6057957 had only a 74.4% match overall (Fig. 1).

In silico serotyping, virulence genes, and MLST profiles

The in silico serotyping of *E. coli* strains revealed that six of the isolates encoded the wzx_{O26} and $fliC_{H11}$ genes. Bovine *E. coli* strain SRR6057957, which was not serotyped by slide agglutination test, encoded wzx_{O182} and $fliC_{H25}$ genes (*E. coli* O182:H25).

As detected by PCR and WGS, both human associated *E. coli* strains, SRR6293238 and SRR6293256, and two of the bovine strains SRR6057923 and SRR5876423, encoded the stx1 gene but not the stx2 gene.

All strains analyzed in this study encoded the following virulence factors: eae, ehxA, espA, espP, gad, iss, lpfA, nleA, nleB, nleC, and tir (Table 1). In comparison to other isolates, the four STEC strains have also encoded additional virulence factors to include celB, efa1, katP, and tccP. Interestingly, toxB was encoded by all the STEC strains except bovine STEC strain SRR5876423 (Table 1).

For in silico MLST typing, all four stx1 positive *E. coli* O26:H11 (2 from human and 2 from cattle) were ST21 while two of the stx1 negative *E. coli* O26:H11 bovine strains were ST29 and the bovine *E. coli* O182:H25 belonged to ST300.

Strain phylogenetic relatedness

The pairwise SNP matrix table using high-quality SNP analysis (Table 2) revealed that both human strains, SRR6293238 and SRR6293256, as well as bovine strain SRR6057923 were identical. Bovine strain SRR5876423 differed from these identical strains by 7 hq-SNP. Bovine strains SRR6057924 and SRR5876420 differed by only 1 hq-SNP from each other but by ≥2000 hq-SNP from the four STEC O26 strains. Bovine *E. coli* strain SRR6057057 differed by ≥30,000 hq-SNP from the other *E. coli* strains (Table 2).

Based on the pairwise SNP matrix data, a distance matrix was calculated using Manhattan coefficient method. An output dendrogram in Newick format was generated with a correlation coefficient of 0.99 (Fig. 2). All four ST 21 STEC O26 strains clustered into one group, while both *E. coli* ST29 strains clustered into a different group, and the *E. coli* O182:H25 was distant from both of these clustered groups (Fig. 2).

Discussion

Previous studies have shown that WGS holds great potential for in-depth characterization of non-O157 STEC strains [12,15,18,20,21]. This potential includes not only serotype identification, but also the detection of virulence factor genes and the ability to distinguish strain relatedness. In this study, we applied WGS to characterize a zoonotic transmission of STEC O26 from cattle to humans.

Of the seven *E. coli* strains analyzed in this study from cattle and humans with diarrhea, bovine *E. coli* strain SRR6057957 could not

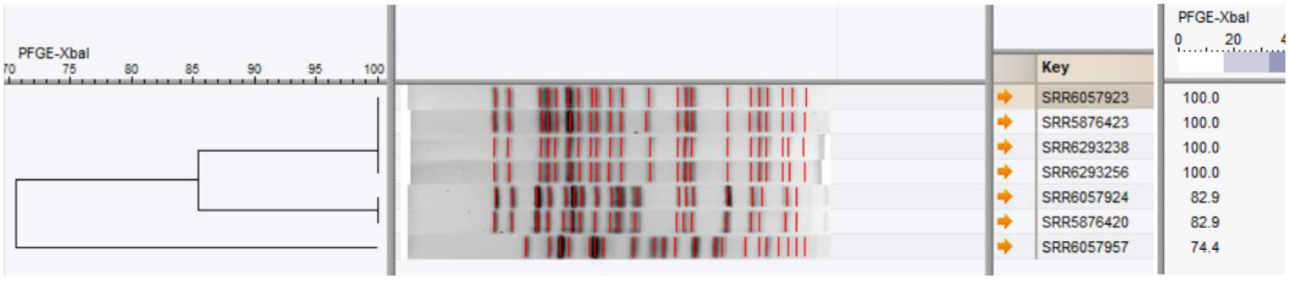

Fig. 1. PFGE fingerprint analysis for the seven isolates represented in this study.
The results showed that isolates represented in the first 4 lanes were indistinguishable from each other while isolates in lanes 5 and 6 were indistinguishable from each other but different from those in the first four lanes. The isolate in lane 7 was distinct from all of the other 6 isolates. BioNumerics 6.6 was used to analyze and interpret band profiling. Calculated strain relatedness, presented as percentage of similarity, was used to construct the phylogentic tree.

Table 1
Sequence types and virulence genes detected by WGS in *E. coli* strains used in this study. +: detected, –: not detected.

	SRR6293238 human	SRR6293256 human	SRR6057923 cow	SRR5876423 cow	SRR6057924 cow	SRR5876420 cow	SRR6057957 cow
MLST sequence type	ST21	ST21	ST21	ST21	ST29	ST29	ST300
astA	+	+	+	+	+	+	–
cba	+	+	+	+	+	+	–
cdtB	–	–	–	–	–	–	+
celB	+	+	+	+	–	–	–
cif	+	+	+	+	+	+	–
eae	+	+	+	+	+	+	+
efa1	+	+	+	+	–	–	–
ehxA	+	+	+	+	+	+	+
espA	+	+	+	+	+	+	+
espB	+	+	+	+	+	+	–
espF	+	+	+	+	+	+	+
espJ	+	+	+	+	+	+	–
espP	+	+	+	+	+	+	+
gad	+	+	+	+	+	+	+
iha	+	+	+	+	+	+	–
iss	+	+	+	+	+	+	+
katP	+	+	+	+	–	–	–
lpfA	+	+	+	+	+	+	+
nleA	+	+	+	+	+	+	+
nleB	+	+	+	+	+	+	+
nleC	+	+	+	+	+	+	+
stx1	+	+	+	+	–	–	–
tccP	+	+	+	+	–	–	–
tir	+	+	+	+	+	+	+
toxB	+	+	+	–	+	+	–

Table 2
Pairwise SNP matrix table showing hqSNPs difference between *E. coli* strains used in this study. Data was generated by lyve-SET (v.2.0.1).

Strain	Source	SRR5876420	SRR5876423	SRR6057923	SRR6057924	SRR6057957	SRR6293238	SRR6293256
SRR5876420	Cow	0	2203	2228	1	30791	2228	2224
SRR5876423	Cow	2203	0	7	2343	32566	7	7
SRR6057923	Cow	2228	7	0	2384	33152	0	0
SRR6057924	Cow	1	2343	2384	0	31487	2378	2380
SRR6057957	Cow	30791	32566	33152	31487	0	33199	33192
SRR6293238	Human	2228	7	0	2378	33199	0	0
SRR6293256	Human	2224	7	0	2380	33192	0	0

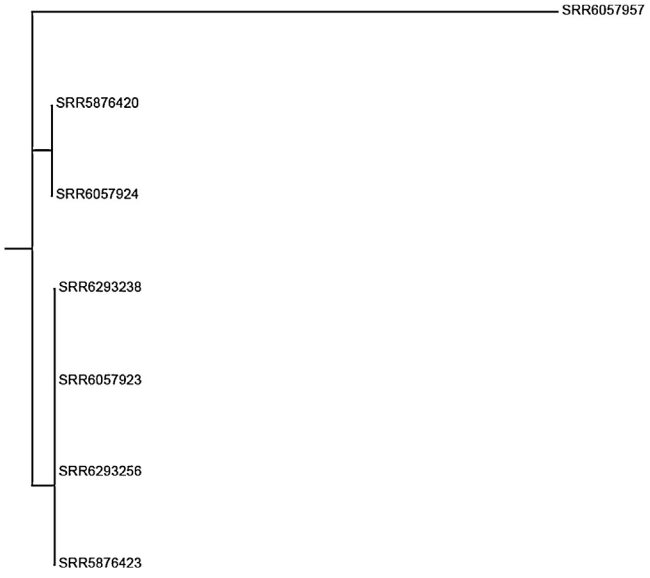

Fig. 2. Phylogenetic analysis of *E. coli* strains analyzed in this study by hq-SNP analysis using Lyve-SET pipeline. The Manhattan coefficient and UPGMA methods were used to compare sets of variables and create phylogenetic tree, respectively. Output dendrogram is presented in Newick format.

be serotyped by the latex agglutination method due to unavailability of serotype O182 antibodies. WGS however was able to provide the serotype. Most public health laboratories serotype *E. coli* strains using the common O groups. However, novel serotypes, cross reactivity, or the unavailability of serotype antibodies require additional testing [14,18,20]. Although previously shown by other reports, this study confirmed that phenotypically untypable isolates could be serotyped using genomic data provided by WGS [12,14,18,20].

PFGE data showed that human STEC O26 strains SRR6293238 and SRR6293256 and STEC bovine strains SRR6057923 and SRR5876423 were indistinguishable from each other (Fig. 1). PFGE did not have the ability to discriminate between the two possible bovine sources of STEC O26. WGS revealed that bovine strain SRR6057923 had an identical SNP matrix to the human strains and differed by only 7 hqSNP from the other bovine strain SRR5876423 suggesting that SRR6057923 strain may have been the origin of human infection. To further support this conclusion, the virulence factor profiles of the human strains SRR6293238 and SRR6293256 were characterized and shown to be identical to that of bovine STEC strain SRR6057923. WGS also indicated that STEC bovine strain SRR5876423 differed by only one virulence factor (toxB) from the human STEC strains (Table 1). Not only was PFGE not able to discriminate STEC O26 strains on a 7 hqSNP level, but it was also not capable of providing any virulence factor profile information, both of which were possible with WGS.

MLST performed on the whole genome sequence also showed that the two human and two bovine STEC O26 strains were characterized as ST21. MLST sequence types for both stx1 negative *E. coli* O26:H11 strains were identified as ST29 and the O182:H25 strain

as ST300. As with PFGE, MLST would have linked the two bovine STEC O26 strains to the human cases, one which was subsequently shown by WGS to not be related. Therefore, WGS has superior discriminatory power compared to MLST.

Several studies support our study finding that WGS has better resolution and more discriminatory power than serotyping, PFGE, and MLST [6,12,15,18,21,25–27]. A 2011 STEC outbreak in Germany representing more than 4000 cases of bloody diarrhea (with 850 cases of HUS) resulting in 50 cases of death were evaluated using serotyping, PFGE, MLST, and other epidemiological tools [25,26]. Using WGS, the *E. coli* O104:H4 isolate responsible for the outbreak was determined to belong to an enteroaggregative *E. coli* strain which had acquired the shigatoxin gene and other antimicrobial resistance determinant genes [25]. This finding was surprising since EHEC O104:H4 is not generally known to be associated with such high morbidity and mortality [26,27]. Other reports utilizing WGS have shown that this technology was able to discriminate organisms causing outbreaks that appeared indistinguishable by routine genotyping assays such as STEC O157, methicillin-resistant Staphylococcus aureus, cholera, and tuberculosis [28–31].

In this study, WGS also provided a comprehensive profile of virulence factors for each strain which helped to obtain a better understanding of the human disease pathogenicity. All four STEC O26 strains in the present study encoded stx1 but not stx2. Studies have shown that strains encoding stx2 are found to be highly virulent and commonly associated with more severe diseases to include HUS than those encoding stx1 [1,2,6,7]. The virulence profiles of the human *E. coli* O26:H11 strains in this study correlated with the clinical picture as neither of the children developed severe infection or HUS.

Based on WGS data, the *E. coli* O26 strains analyzed in this study could be categorized into two groups. The first group, EPEC group, contained the SRR6057924 and SRR5876420 strains which were stx negative and eae positive. The second group, referred to as the EHEC group, was composed of both human strains, SRR6293238 and SRR6293256, and the bovine strains SRR6057923 and SRR5876423. This group of strains encoded both the stx1 and eae genes. This characterization by WGS allowed for a better classification of organisms during an outbreak.

Epidemiological studies using serological or other molecular methods such as PFGE and/or MLST limit the ability of epidemiologists and microbiologists to compare data between different outbreaks due to the varied discriminatory power of each method [12,15,18,20]. These traditional methods may even misclassify outbreak related organisms, which can hinder further epidemiological investigations [31]. WGS has an advantage over these techniques in that it provides information on the whole genome of the organism using a single technology. This information can be then analyzed based on requested serotyping, sequencing types, antimicrobial resistance determinants, plasmid profiles, and/or virulence factor profiling. Although WGS requires an expensive instrument, bioinformatics knowledge, large computational servers, data standardization, and global guidelines, several studies showed that WGS is favored over serological and other molecular epidemiological methods when compared based upon cost, hands-on time, and technique optimization [15,18,20,21,31,32].

The application of WGS has enhanced laboratory testing by reducing costs and lowering the turn-around-times while increasing the amount of information generated following the availability of complete genome sequencing. WGS was able to provide characterization of the STEC O26 strains in a one-step method within about 4 days while growth-based traditional methods required several weeks. These tests included species identification, agglutination assays for serotyping, and assays for antimicrobial susceptibility testing. In addition, several PCR assays for virulence factors and for MLST determination as well as PFGE fingerprinting and Sanger sequencing assays were also needed for strain characterization. All of these methods require comprehensive workflows, expertise, expensive reagents, and daily quality control measures. As a means to better expand on the capabilities of the PulseNet program, the Centers for Disease Control and Prevention has replaced PFGE and other traditional methods with WGS [33,34].

In conclusion, this study showed that WGS is a robust, reliable, one-step process with high discriminatory power in the analysis of a zoonotic transmission of STEC O26 from cattle to humans. This method was shown to have the capability to provide rapid data for identification, serotyping, sequence typing, and virulence profiling of a non-O157 pathogen. The results of this study also confirmed that WGS is superior to traditional epidemiological methods in outbreak investigations of STEC non-O157. Characterization of virulence factors and other genetic features using WGS is expected to provide a better understanding of bacterial pathogenicity in the future.

Author contributions

All authors have contributed to the manuscript. BA: study design and writing and editing the manuscript; ELM: technical support and writing the section of materials and methods; ACB and ZW: analysis and interpretation of bioinformatics data. SHH and PCI: study design and proof reading of the manuscript.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Funding information

The authors would like to thank CDC for partially supporting this study by a grant # CK14-140105PPHF18.

Ethical approval

Experiments were performed only on bacterial colonies growing on agar plates. No experiments were conducted on human subjects. Therefore, no IRB approval was needed.

References

[1] Delannoy S, Mariani-Kurkdjian P, Bonacorsi S, Liguori S, Fach P. Characteristics of emerging human-pathogenic *Escherichia coli* O26:H11 strains isolated in France between 2010 and 2013 and carrying the stx2d gene only. J Clin Microbiol 2015;53(2):486–92.
[2] Bielaszewska M, Mellmann A, Bletz S, Zhang W, Kock R, Kossow A, et al. Enterohemorrhagic *Escherichia coli* O26:H11/H-: a new virulent clone emerges in Europe. Clin Infect Dis 2013;56(10):1373–81.
[3] Yeni F, Yavas S, Alpas H, Soyer Y. Most common foodborne pathogens and mycotoxins on fresh produce: a review of recent outbreaks. Crit Rev Food Sci Nutr 2016;56(9):1532–44.
[4] Thomas MK, Murray R, Flockhart L, Pintar K, Fazil A, Nesbitt A, et al. Estimates of foodborne illness-related hospitalizations and deaths in Canada for 30 specified pathogens and unspecified agents. Foodborne Pathog Dis 2015;12(10):820–7.
[5] Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, et al. Foodborne illness acquired in the United States–major pathogens. Emerg Infect Dis 2011;17(1):7–15.
[6] Amezquita-Lopez BA, Soto-Beltran M, Lee BG, Yambao JC, Quinones B. Isolation, genotyping and antimicrobial resistance of Shiga toxin-producing *Escherichia coli*. J Microbiol Immunol Infect 2018;51(4):425–34.
[7] Etcheverria AI, Padola NL. Shiga toxin-producing *Escherichia coli*: factors involved in virulence and cattle colonization. Virulence 2013;4(5):366–72.
[8] Bosilevac JM, Koohmaraie M. Prevalence and characterization of non-O157 shiga toxin-producing *Escherichia coli* isolates from commercial ground beef in the United States. Appl Environ Microbiol 2011;77(6):2103–12.
[9] Mellor GE, Fegan N, Duffy LL, Mc MK, Jordan D, Barlow RS. National survey of shiga toxin-producing *Escherichia coli* serotypes O26, O45, O103, O111, O121, O145, and O157 in Australian beef cattle feces. J Food Prot 2016;79(11):1868–74.

[10] Bletz S, Bielaszewska M, Leopold SR, Kock R, Witten A, Schuldes J, et al. Evolution of enterohemorrhagic *Escherichia coli* O26 based on single-nucleotide polymorphisms. Genome Biol Evol 2013;5(10):1807–16.
[11] Boerlin P, McEwen SA, Boerlin-Petzold F, Wilson JB, Johnson RP, Gyles CL. Associations between virulence factors of shiga toxin-producing *Escherichia coli* and disease in humans. J Clin Microbiol 1999;37(3):497–503.
[12] Gonzalez-Escalona N, Toro M, Rump LV, Cao G, Nagaraja TG, Meng J. Virulence gene profiles and clonal relationships of *Escherichia coli* O26:H11 isolates from feedlot cattle as determined by whole-genome sequencing. Appl Environ Microbiol 2016;82(13):3900–12.
[13] Smith AM, Bosco KJ, Nicol MP, Kleynhans J, McCulloch M, Duze ST, et al. Genome sequence for shiga toxin-producing *Escherichia coli* O26:H11, associated with a cluster of hemolytic-uremic syndrome cases in South Africa, 2017. Genome Announc 2017;5(38).
[14] Monaghan A, Byrne B, Fanning S, Sweeney T, McDowell D, Bolton DJ. Serotypes and virulence profiles of non-O157 Shiga toxin-producing *Escherichia coli* isolates from bovine farms. Appl Environ Microbiol 2011;77(24):8662–8.
[15] Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, et al. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol 2014;52(5):1501–10.
[16] Jenkins C, Evans J, Chart H, Willshaw GA, Frankel G. *Escherichia coli* serogroup O26—a new look at an old adversary. J Appl Microbiol 2008;104(1):14–25.
[17] Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, et al. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol 2012;50(4):1355–61.
[18] Dallman TJ, Byrne L, Launders N, Glen K, Grant KA, Jenkins C. The utility and public health implications of PCR and whole genome sequencing for the detection and investigation of an outbreak of shiga toxin-producing *Escherichia coli* serogroup O26:H11. Epidemiol Infect 2015;143(8):1672–80.
[19] Carter JM, Lin A, Clotilde L, Lesho M. Rapid, multiplexed characterization of shiga toxin-producing *Escherichia coli* (STEC) isolates using suspension array technology. Front Microbiol 2016;7:439.
[20] Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and easy in silico serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. J Clin Microbiol 2015;53(8):2410–26.
[21] Chattaway MA, Dallman TJ, Gentle A, Wright MJ, Long SE, Ashton PM, et al. Whole genome sequencing for public health surveillance of shiga toxin-producing *Escherichia coli* other than serogroup O157. Front Microbiol 2016;7(258):251–4.
[22] Ribot EM, Fair MA, Gautom R, Cameron DN, Hunter SB, Swaminathan B, et al. Standardization of pulsed-field gel electrophoresis protocols for the subtyping of *Escherichia coli* O157:H7, Salmonella, and Shigella for PulseNet. Foodborne Pathog Dis 2006;3(1):59–67.
[23] Katz LS, Griswold T, Williams-Newkirk AJ, Wagner D, Petkau A, Sieffert C, et al. A comparative analysis of the Lyve-SET phylogenomics pipeline for genomic epidemiology of foodborne pathogens. Front Microbiol 2017;8(375).
[24] Fratamico PM, Sackitey SK, Wiedmann M, Deng MY. Detection of *Escherichia coli* O157:H7 by multiplex PCR. J Clin Microbiol 1995;33(8):2188–91.
[25] Rasko DA, Webster DR, Sahl JW, Bashir A, Boisen N, Scheutz F, et al. Origins of the *E. Coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. N Engl J Med 2011;365(8):709–17.
[26] Gault G, Weill FX, Mariani-Kurkdjian P, Jourdan-da Silva N, King L, Aldabe B, et al. Outbreak of haemolytic uraemic syndrome and bloody diarrhoea due to *Escherichia coli* O104:H4, south-west France, June 2011. Euro Surveill 2011;16(26).
[27] Mariani-Kurkdjian P, Bingen E, Gault G, Jourdan-Da Silva N, Weill FX. *Escherichia coli* O104:H4 south-west France, June 2011. Lancet Infect Dis 2011;11(10):732–3.
[28] Chin CS, Sorenson J, Harris JB, Robins WP, Charles RC, Jean-Charles RR, et al. The origin of the Haitian cholera outbreak strain. N Engl J Med 2011;364(1):33–42.
[29] Gardy JL, Johnston JC, Ho Sui SJ, Cook VJ, Shah L, Brodkin E, et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. N Engl J Med 2011;364(8):730–9.
[30] Koser CU, Holden MT, Ellington MJ, Cartwright EJ, Brown NM, Ogilvy-Stuart AL, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. N Engl J Med 2012;366(24):2267–75.
[31] Rumore J, Tschetter L, Kearney A, Kandar R, McCormick R, Walker M, et al. Evaluation of whole-genome sequencing for outbreak detection of verotoxigenic *Escherichia coli* O157:H7 from the Canadian perspective. BMC Genomics 2018;19(1):870.
[32] Holmes A, Dallman TJ, Shabaan S, Hanson M, Allison L. Validation of whole-genome sequencing for identification and characterization of shiga toxin-producing *Escherichia coli* to produce standardized data to enable data sharing. J Clin Microbiol 2018;56(3):1–10.
[33] Lindsey RL, Pouseele H, Chen JC, Strockbine NA, Carleton HA. Implementation of whole genome sequencing (WGS) for identification and characterization of shiga toxin-producing *Escherichia coli* (STEC) in the United States. Front Microbiol 2016;7(766).
[34] Jackson BR, Tarr C, Strain E, Jackson KA, Conrad A, Carleton H, et al. Implementation of Nationwide Real-time Whole-genome Sequencing to Enhance Listeriosis Outbreak Detection and Investigation. Clin Infect Dis 2016;63(3):380–6.