Virulence gene profiles and phylogeny of Shiga toxin-positive Escherichia coli strains isolated from FDA regulated foods during 2010-2017

Narjol González-Escalona, Julie Ann Kase

Division of Microbiology, Center for Food Safety and Applied Nutrition, Food and Drug Administration, College Park, MD, United States of America

Abstract

Illnesses caused by Shiga toxin-producing Escherichia coli (STECs) can be life threatening, such as hemolytic uremic syndrome (HUS). The STECs most frequently identified by USDA’s Microbiological Data Program (MDP) carried toxin gene subtypes stx1a and/or stx2a. Here we described the genome sequences of 331 STECs isolated from foods regulated by the FDA 2010–2017, and determined their genomic identity, serotype, sequence type, virulence potential, and prevalence of antimicrobial resistance. Isolates were selected from the MDP archive, routine food testing by FDA field labs (ORA), and food testing by a contract company. Only 276 (83%) strains were confirmed as STECs by in silico analysis. Foods from which STECs were recovered included cilantro (6%), spinach (25%), lettuce (11%), and flour (9%). Phylogenetic analysis using core genome MLST revealed these STEC genomes were highly variable, with some clustering associated with ST types and serotypes. We detected 95 different sequence types (ST); several ST were previously associated with HUS: ST21 and ST29 (O26:H11), ST11 (O157:H7), ST33 (O91:H14), ST17 (O103:H2), and ST16 (O111:H-). in silico virulome analyses showed ~ 51% of these strains were potentially pathogenic [besides stx gene they also carried eae (25%) or 26% saa (26%)]. Virulence gene prevalence was also determined: stx1 only (19%); stx2 only (66%); and stx1/stx2 (15%). Our data form a new WGS dataset that can be used to support food safety investigations and monitor the recurrence/emergence of E. coli in foods.

Introduction

Shiga toxin-producing Escherichia coli (STECs) have the potential to cause infections, from mild to life-threatening outcomes such as hemolytic uremic syndrome (HUS). STECs causing HUS, hemorrhagic colitis and bloody diarrhea are known as enterohemorrhagic E. coli (EHEC). Among the most common EHECs are O157:H7, O26, O121, O103, O111, and O145. O157:H7 strains are responsible for most foodborne outbreaks in the last two decades [1].
while non-O157 serogroups, O26, O121, O103, O111, and O145 are the second most common cause of EHEC foodborne infections in the US [2,3] and worldwide [4–7]. Each year in the US, O157:H7 causes approximately 95,000 cases with 2,150 hospitalizations, while non-O157 STECs are responsible for an estimated 170,000 cases [3]. These serotypes carry Shiga toxin genes (stx1 and/or stx2) and there are at least 130 EHEC serotypes that have been recovered from human patients. In 2011 the US Department of Agriculture Food Safety and Inspection Services (USDA FSIS), declared O26 and five other non-O157 serogroups, O45, O103, O111, O121, and O145 as adulterants in ground beef and non-intact beef products, and in mid-2012 began testing for these pathogens in both domestic and imported beef trimmings [8].

In order to cause illness, STEC strains need a set of genes that allow them to attach, colonize, and produce and secrete Shiga toxin protein [9–12]. STECs have the capacity to produce attaching and effacing (A/E) lesions on intestinal mucosa, mediated by proteins found on the locus of enterocyte effacement (LEE) Pathogenicity Island. These include EAE protein (intimin, coded by the eae gene), secreted effector proteins (Esp), a TIR receptor, and other T3SS effectors present in the LEE island. However, some STECs do not carry eae (LEE-negative STEC strains). They possess other genes believed to compensate for the lack of eae or LEE island (e.g. saa) and still caused sporadic cases of HUS [13,14]. Other putative virulence genes (ehxA, espP, etpD, toxB, katP, subA, saa, and sab genes, among others) are usually located in a plasmid referred to as the virulence plasmid to differentiate it from other possible plasmids that can be carried by the same strain [9,11,15]. All STECs have a virulence plasmid, although differing in sizes and gene content (either virulence or plasmid maintenance genes). Although the precise role of ehxA in STEC pathogenesis remains to be elucidated, several studies indicate an association of ehxA in clinical disease since 1) ehxA was found to be produced by many STEC associated with diarrheal disease and HUS [16–18], and 2) serum samples from HUS patients have been shown to react specifically to ehxA [19].

STECs can be transmitted by various means with food remaining the predominant transmission route [1]. Among the illnesses caused by STECs in FDA regulated food products (FRFDA), fresh produce has been implicated in several outbreaks, as well as some other atypical commodities, such as flour [20]. Leafy greens and other agricultural food crops are particularly susceptible to contamination since they are grown in close contact with the ground where runoff from livestock areas, particularly cattle, contaminated irrigation water, manure used as fertilizer, and the intrusion of wildlife into growing fields can occur [21]. Many of these same items are consumed raw and possibly with little cleaning. Noteworthy E. coli outbreaks reported by the Center of Disease and Control (CDC) in the US in the last 10 years are: in 2009—beef (O157:H7) and prepackaged cookie dough (O157:H7); in 2010—cheese (O157:H7), romaine lettuce (O145) and beef (O157:H7); in 2011—romaine lettuce (O157:H7), Lebanon bologna (O157:H7), and in-shell hazelnuts (O157:H7); in 2012—spinach and spring mix blend (O157:H7), unknown source (O145), and raw clover sprouts (O26); in 2013—ready-to-eat salads (O157:H7), and frozen food products (O121); in 2014—raw clover sprouts (O121), and ground beef (O157:H7); in 2015—rotisserie chicken salad (O157:H7), and Mexican-style restaurant chain (O26); in 2016—flour (O121 and O26), and alfalfa sprouts (O157); 2017—leafy greens (O157:H7), and soy nut butter (O157:H7); and in 2018—there has been an outbreak linked to romaine lettuce caused by O157: H7 (https://www.cdc.gov/ecoli/outbreaks.html). Beyond the noted outbreaks, there have been several reports on STECs found in FRFDA [22,23]. The most comprehensive survey was the USDA Microbiological Data Program (MDP) that collected domestic and imported fresh fruit and vegetable samples from primarily terminal markets and wholesale distribution centers from 2001–2012 (https://www.ams.usda.gov/datasets/mdp/mdp-program-data-and-reports). This program tested approximately 15,000 samples annually, and tested for the presence of Salmonella, E. coli O157:H7, and other
STECs. STEC were most frequently found in spinach samples (0.5%), and of the 132 isolated STECs, 9% were found to carry eae. The most prevalent Shiga toxin variants found were stx1a (22%) and/or stx2a (56%) [23]. However, little other information about the genome content of those strains is publicly available.

Whole genome sequencing (WGS) technology is reshaping food safety and food-borne illness investigations [24]. The use of WGS is becoming more useful as the cost of bacterial genome sequencing decreases every year. WGS’s cost per bacterial sequence is now comparable to PFGE. There are many attractive attributes with regards to the use of WGS in analyzing food samples including the potential to identify all pathogens present in that sample [25]. Other applications of WGS may include assistance in the following: identifying genes that allow for resistance/survival or virulence of certain bacterial strains [26–28], establishing phylogenetic relationships among old strains of STECs isolated from either clinical cases or environmental samples [7,29,30], and identifying matches between environmental and outbreak strains during outbreak scenarios [26,30–32]. Furthermore, using WGS can help in identifying matches among bacterial strains isolated from environmental samples in production facilities and locating contamination sources [33]. It can also be extremely helpful in establishing a mechanism of evolution among pathogens [34]. For example, the 2011 outbreak in Germany was linked to fenugreek seeds caused by an E. coli strain with a genomic backbone and virulence traits of entero-aggregative E. coli (EAEC) but had acquired a stx phage (stx2a gene variant) and caused a more aggressive disease with high HUS rate cases [31,35]. This event highlighted the high plasticity of the E. coli genomes and it constitutes a warning of the possible rise of new “hybrid pathotype” strains.

Therefore, we wanted to further characterize and catalog historical strains of STECs isolated from FRFDA by performing WGS analysis of every STEC strain isolated by the MDP and other FDA surveillance programs, as well as some FDA historical isolates. Most STEC strains were isolated during active surveillance programs and were not outbreak related with the exception of some strains isolated during a flour outbreak of 2016 (ST155 –O121:H19). This work establishes the first genomic dataset of FRFDA STEC strains, which in turn, will allow improved surveillance for both recurrence and the emergence of new strains that may impact our food supply. A total of 296 presumptive STECs were isolated during 2010–2017, and 35 additional STECs were historical isolates from our collection. The 331 presumptive STEC strains were analyzed for virulence genes [encompassing all E. coli virulent types—STEC, entero-pathogenic E. coli (EPEC), entero-toxigenic E. coli (ETEC), entero-invasive E. coli (EIEC), and EAEC], in silico MLST, and antibiotic resistance genes. Finally, their phylogenetic relationships and diversity were determined by whole genome phylogeny analysis using an allele-based whole genome multilocus sequence analysis (MLST) or core genome MLST analysis (cgMLST).

**Materials and methods**

**Bacterial strains and media**

_E. coli_ (n = 331) presumptive Shiga toxin-positive strains used in this study are listed in S1 Table. Each strain was assigned a CFSAN number for future tracking. The FRFDA strains were isolated by MDP program (n = 196), FDA Office of Regulatory Affairs (ORA) laboratories (n = 74) during food surveillance testing, and a contracting lab (n = 63) during 2012–2017 in the US.

**DNA preparation**

Genomic DNA from each strain was isolated from overnight cultures using the DNeasy Blood and Tissue Kit (QIAGEN, Valencia, CA), following the manufacturer’s instructions. The
resultant DNA extract was stored at -20˚C until used as a template for whole genome sequencing. The concentration was determined using a Qubit double-stranded DNA HS assay kit and a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA), according to the manufacturer’s instructions.

**Whole genome sequencing, contig assembly and annotation**

The genomes of the strains sequenced in our laboratory used an Illumina MiSeq sequencer (Illumina, San Diego, CA), with the 2x250 bp pair-end chemistry according to the manufacturer’s instructions, at approximately 80X average coverage. The genome libraries were constructed using the Nextera XT DNA sample prep kit (Illumina). Genomic sequence contigs were *de novo* assembled using default settings within CLC Genomics Workbench v9.5.2 (QIAGEN) with a minimum contig size threshold of 500 bp in length.

**in silico MLST phylogenetic analysis**

The initial analysis and identification of the strains were performed using an *in silico* E. coli MLST approach, based on the information available at the E. coli MLST website (http://mlst.warwick.ac.uk/mlst/dbs/Ecoli) and using Ridom SeqSphere+ software v2.4.0 (Ridom; Münster, Germany) (http://www.ridom.com/seqsphere). Seven housekeeping genes (*dnaE*, *gyrB*, *recA*, *ddS*, *pntA*, *pyrC*, and *tnaA*), described previously for E. coli [36], were used for MLST analysis. The same E. coli MLST database was also used to assign numbers for alleles and STs.

**in silico serotyping, determination of virulence genes, and antimicrobial resistance genes identification**

We used Ridom software for performing batch screening of determinants for each of the sequenced genomes.

The serotype of each strain analyzed in this study was confirmed using the genes deposited in the Center for Genomic Epidemiology (http://www.genomicepidemiology.org) for E. coli as part of their web-based serotyping tool (SerotypeFinder 1.1 - https://cge.cbs.dtu.dk/services/SerotypeFinder) [37,37,38]. Each whole genome sequence was screened for O-type or H-type genes. The virulence genes present in each strain were determined using the genes deposited in the Center for Genomic Epidemiology (http://www.genomicepidemiology.org) for E. coli as part of their VirulenceFinder 1.5 web-based tool (https://cge.cbs.dtu.dk/services/VirulenceFinder) [38]. Each WGS was screened for the presence of 102 virulence genes (95 virulence genes previously reported here [27] plus 7 additional genes) (S2 Table). These 102 virulence genes included a repertoire representing genes found in different E. coli pathotypes (ETEC, STEC, EAEC, and EPEC) in order to detect any possible E. coli hybrid present. Antimicrobial resistance genes present in each sequenced genomes were identified by using the genes deposited in the Center for Genomic Epidemiology (http://www.genomicepidemiology.org) as part of their Resfinder 2.1 web-based tool (https://cge.cbs.dtu.dk/services/ResFinder) [39]. Each WGS was screened for the presence of those AMR genes reported in the database.

**Phylogenetic relationship of the strains by cgMLST analysis**

The phylogenetic relationship of the strains was assessed by a core genome multilocus sequence typing (cgMLST) analysis using Ridom SeqSphere+ software v2.4.0. The genome of O157:H7 strain Sakai (NC_002695.1) was used as the reference for the cgMLST. This E. coli strain has 5,204 genes, of which 3,860 genes (core genes) were present in the six genomes used as comparison to generate the cgMLST scheme (NC_011353.1 –O157:H7
strain EC4115, NC_002655.2—O157:H7 strain EDL933, NC_013008.1 —O157:H7 strain TW14359, NC_013941.1—O55:H7 strain CB9615, NC_017656.1—O55:H7 strain RM12579, and NC_017906.1 —O157:H7 strain Xuzhou21). While 791 genes were found in some of the compared genomes, the remaining of the genes were eliminated from the analysis since they were paralogous or pseudogenes. Therefore, a total of 4,651 genes were used as templates for the analysis of the STECs from this study. After eliminating loci that were missing from the genome of any strain used in our analyses, we performed a cgMLST analysis. These remaining loci were considered the core genome shared by the analyzed strains. We used Nei’s DNA distance method [40] for calculating the matrix of genetic distance, taking into consideration only the number of same/different alleles in the core genes. A Neighbor-Joining (NJ) tree using the appropriate genetic distances was built after the cgMLST analysis. The discriminatory index was calculated with the Ridom software using the Simpson’s discriminatory index as described [41]; cgMLST uses the alleles number of each loci for determining the genetic distance and builds the phylogenetic tree. The use of allele numbers reduces the influence of recombination in the dataset studied and allows for fast clustering determination of genomes.

Nucleotide sequence accession numbers
The draft genome sequences of 331 E. coli strains used in our study are available in GenBank under the accession numbers listed in S1 Table.

Results
Presence of STEC in FDA regulated foods
Among 331 suspected STEC strains isolated from FRFDA between 2003–2017 and sequenced by several labs and deposited at NCBI, only 276 were confirmed to be STECs by in silico analysis for the presence of either stx1 or stx2 (S1 Table). Those that were stx negatives either might have lost their phages or were initially misidentified. Of the 196 identified by MDP and sequenced by our lab, 92% carried either stx1 or 2 (181/196). Of the 74 strains which genomes were retrieved from NCBI and were initially isolated and sequenced by FDA ORA, 94% carried either stx1 or 2 (70/74). Of the 63 E. coli strains isolated and sequenced by a FDA contracting laboratory, 43% carried either stx1 or 2 (25/61). The frequency of isolation of STECs from foods are listed in Table 1. STECs were isolated from 22 food commodities. Most of these STECs were isolated from spinach (32%), flour (21%), lettuce (13%), and cilantro (12%).

The frequency of STEC isolation per year, their sequence type, food commodity and state of isolation (if available) is listed in Table 2. The minimum number of STECs recovered from FDA regulated foods per year during the period 2010–2017 was 12 in 2012 and the maximum was 53 in 2016, for a median of 30 STECs per year. We used 2010 as our starting year, since the number of strains prior to that year were found sporadically and came from our STEC historical collection. The STEC analyzed strains were isolated in 22 states.

Characterization of STEC strains by in silico MLST
Among the 276 STECs analyzed in this study we identified 95 different sequence types (STs) by in silico MLST. Strains belonging to a ST were isolated between 1 to 12 times in the period studied (2010–2017) (Tables 3 and S3). The majority of the STs identified were observed only one time (45%).
Characterization of STEC strains by serotyping, and virulence gene profiles

However, a strain belonging to a known ST that caused hemorrhagic colitis (HC) is not enough to predict the likelihood of it causing disease illness. Therefore, we further characterized these STECs by in silico virulence determination and their predicted serotype (Table 4). The detailed in silico analysis for presence of virulence genes and serotype is listed in S2 Table. Table 4 lists only the serotype and some of the most known virulence genes for each strain: stx1 variants, stx2 variants, eae variants, ehxA, espP, etpD, toxB, katP, subA, saa, and sab. We identified at least 81 different serotypes among the 276 STECs sequenced (Table 4). Many of the O types were not present in our O types database and were listed as unknown or the wzx and wzy gene were not assembled. Among those some of the most common clinical STECs serotypes were identified, such as: O157:H7, O26:H11, O113:H21, O121:H19, O91:H21, O103:H2, and O111:H8. Other serotypes like O153/O178:H19 could not be discerned by in silico analysis since both O types have identical wzx and wzy gene sequences. This issue is discussed in more detail elsewhere [37].

Adherence factors eae and saa genes were found in 67 (24%) and 72 (26%), respectively (Table 4). Shiga toxin genes were present as follows: stx1- 53 (19%) (variants a and c), stx2-184 (67%) (variants a, b, c, d, d/e, e, and g), while stx1+stx2–39 (15%). Among the 184 STECs carrying only stx2 genes, 144 of them carried variants 2a, 2d, or 2e; while the remaining 40 carried variants 2e, d/e, or g. Other putative virulence genes were found as follows: exhA gene was

Table 1. Frequency of STEC isolated by food commodity.

| Commodity                  | No. strains | %   |
|----------------------------|-------------|-----|
| Spinach                    | 88          | 31.88 |
| Flour                      | 59          | 21.38 |
| Lettuce                    | 35          | 12.68 |
| Cilantro (coriander)       | 32          | 11.59 |
| Cheese                     | 9           | 3.26  |
| Leafy greens               | 8           | 2.90  |
| Kale                       | 7           | 2.54  |
| Basil                      | 6           | 2.17  |
| Pepper                     | 6           | 2.17  |
| Alfalfa sprouts            | 3           | 1.09  |
| Cantaloupe                 | 3           | 1.09  |
| Parsley                    | 3           | 1.09  |
| Tomatoes                   | 3           | 1.09  |
| Creamy soy Nut butter      | 2           | 0.72  |
| pizza dough dry mix        | 2           | 0.72  |
| sprouts                    | 2           | 0.72  |
| Almond                     | 1           | 0.36  |
| oats animal feed           | 1           | 0.36  |
| Celery                     | 1           | 0.36  |
| Clover sprouts             | 1           | 0.36  |
| cucumbers                  | 1           | 0.36  |
| animal feed                | 1           | 0.36  |
| environmental              | 2           | 0.72  |
| Total                      | 276         | 100.00 |

The stx-negative strains were eliminated from this analysis.

https://doi.org/10.1371/journal.pone.0214620.t001
present in 169 (61%), espP was present in 118 (43%), katP in 24 (9%), etpD in 4 (2%), and finally toxB was present in 10 (4%).

### Presence of antimicrobial resistance genes

Thirty-three of the 276 STEC strains (12%) carried antimicrobial resistance genes (Table 5). Thirty of them carried multiple antibiotic resistance genes while the remaining three carried a single gene (tetA-IEH-NGS-ECO-00231, FDA00011218, and CFSAN051521). Among the antimicrobial classes observed were genes resistant to aminoglycosides, beta-lactamases, macrolides, phenicols, quinolones, sulphonamides, tetracyclines, and trimethoprim.

### Phylogenetic relationship of the STEC strains by cgMLST analysis

The phylogenetic relationships among the 276 STECs from this study determined by cgMLST analysis is shown in Fig 1. The genome of O157:H7 strain Sakai (NC_002695.1) was used as the reference for the cgMLST. The initial phylogenetic analysis [Neighbor-Joining (NJ) tree] based on gene differences (allele based) among these 276 STECs (Fig 1) revealed a complex evolutionary history with the existence of multiple, highly diverse genomic variants of strains isolated from RFFDA. Some of these genomes formed discrete groups and clustering was consistent with their ST (ex. all ST655 strains clustered together). A further analysis by a minimum
Table 3. STs observed and number of strains included in each ST. Additionally information is provided for strains belonging to those STs such as: link to human cases, link to EHEC cases, and known serotypes. These additional reports are based on what it is reported in the *E. coli* section of the Enterobase database (http://enterobase.warwick.ac.uk).

| ST | No. Strains | %   | Human cases | Reported as EHEC<sup>b</sup> | Known serotypes |
|----|-------------|-----|-------------|----------------------------|----------------|
| 655| 38          | 13.67| +           | +                          | O121:H19       |
| 4496| 12     | 4.32 | +           | NR                         | O8:H28         |
| 2008| 12        | 4.32 | +           | NR                         | Ounk:H2/40     |
| 21 | 8          | 2.88 | +           | +                          | O26:H11/-      |
| 297| 9          | 3.24 | +           | - (UPEC, APEC)              | diverse serotypes |
| 205| 7          | 2.52 | +           | - (UPEC)                   | NR             |
| 43 | 5          | 1.80 | +           | - (ETEC, EAEC)             | O6:H10         |
| 154| 5          | 1.80 | +           | - (EPEC, APEC)             | diverse serotypes |
| 173| 5          | 1.80 | +           | NR                         | O181:H49       |
| 295| 5          | 1.80 | +           | - (EAEC, UPEC, ExPEC)      | diverse serotypes |
| 677| 5          | 1.80 | +           | +                          | diverse serotypes |
| 747| 8          | 2.88 | +           | - (ETEC)                   | diverse serotypes |
| 11 | 3          | 1.08 | +           | +                          | O157:H7/-      |
| 16 | 2          | 0.72 | +           | +                          | O111:H8/21/-   |
| 17 | 2          | 0.72 | +           | +                          | O103:H2/-      |
| 25 | 1          | 0.36 | +           | +                          | O128:H2        |
| 29 | 1          | 0.36 | +           | +                          | O26:H11        |
| 32 | 1          | 0.36 | +           | +                          | O145:H-        |
| 33 | 2          | 0.72 | +           | +                          | O91:H14        |
| 119| 1          | 0.36 | +           | +                          | O165:H25/28    |
| 223| 4          | 1.44 | +           | +                          | O84:H2/K+      |
| 306| 2          | 0.72 | +           | +                          | O154/Ounk:H20  |
| 329| 2          | 0.72 | +           | +                          | O15:H16/K+     |
| 657| 1          | 0.36 | +           | +                          | diverse serotypes |
| 675| 1          | 0.36 | +           | +                          | O17:H19        |
| 679| 3          | 1.08 | +           | +                          | O163:H19       |
| 724| 1          | 0.36 | +           | +                          | O154/Ounk:H20  |
| 5299| 2       | 0.72 | +           | NR                         | O8:H49         |
| 325| 1          | 0.36 | +           | NR                         | O15:H16/K+     |
| 6639| 1      | 0.36 | +           | NR                         | O174:H21/36    |
| 661| 2          | 0.72 | +           | NR                         | O174:H2        |
| 662| 2          | 0.72 | +           | NR                         | diverse serotypes |
| 691| 2          | 0.72 | +           | NR                         | diverse serotypes |
| 692| 3          | 1.08 | +           | NR                         | O74:H42        |
| 718| 3          | 1.08 | +           | NR                         | O168:H8        |
| 723| 4          | 1.44 | +           | NR                         | O103:H11       |
| 942| 1          | 0.36 | +           | NR                         | O116:H28       |
| 955| 1          | 0.36 | +           | NR                         | O139:H1/6      |
| 993| 2          | 0.72 | +           | NR                         | O100:H30       |
| 1792| 1        | 0.36 | +           | NR                         | O111:H8        |
| 1817| 1       | 0.36 | +           | NR                         | O104:H7        |
| 1967| 2        | 0.72 | +           | NR                         | O103:H2        |
| 2388| 1       | 0.36 | +           | NR                         | O15           |
| 2520| 3        | 1.08 | +           | NR                         | O116:H49       |
| 3759| 1        | 0.36 | +           | NR                         |              |
| 5973| 2        | 0.72 | +           | NR                         | Ounk:H2        |

(Continued)
Table 3. (Continued)

| ST# | No. Strains | % | Human cases | Reported as EHECb | Known serotypes |
|-----|-------------|---|-------------|------------------|----------------|
| 6475| 2           | 0.72 | +           | NR               | O17/077:H45     |
| 10  | 2           | 0.72 | +           | - (mainly EAEC, UPEC, ETEC) | diverse serotypes |
| 35  | 1           | 0.36 | +           | - (EPEC or UPEC) | O154:H4, O145:H34/31 |
| 40  | 1           | 0.36 | +           | - (EAEC)         | O111ac:H21      |
| 58  | 1           | 0.36 | +           | - (EAEC, UPEC, ExPEC, APEC) | diverse serotypes |
| 88  | 2           | 0.72 | +           | - (EAEC, UPEC, ExPEC, APEC) | diverse serotypes |
| 101 | 3           | 1.08 | +           | - (EAEC, UPEC, ExPEC, APEC) | diverse serotypes |
| 156 | 1           | 0.36 | +           | - (UPEC, ExPEC)  | diverse serotypes |
| 162 | 1           | 0.36 | +           | - (UPEC, APEC)   | O8:H19          |
| 342 | 2           | 0.72 | +           | - (EPEC)         | O177:NM         |
| 394 | 1           | 0.36 | +           | - (EAEC, UPEC)   | diverse serotypes |
| 398 | 1           | 0.36 | +           | - (ExPEC)        | diverse serotypes |
| 442 | 3           | 1.08 | +           | - (EPEC)         | O146:H21        |
| 443 | 2           | 0.72 | +           | - (UPEC)         | NR              |
| 446 | 2           | 0.72 | +           | - (APEC)         | diverse serotypes |
| 515 | 1           | 0.36 | +           | - (EAEC)         | O2:H9           |
| 641 | 1           | 0.36 | +           | - (ExPEC)        | diverse serotypes |
| 642 | 3           | 1.08 | +           | - (EPEC)         | diverse serotypes |
| 706 | 1           | 0.36 | +           | - (UPEC)         | diverse serotypes |
| 906 | 3           | 1.08 | +           | - (UPEC)         | diverse serotypes |
| 1431| 1           | 0.36 | +           | - (ExPEC)        | O8:H19/30       |
| 1727| 4           | 1.44 | +           | - (mostly nonpathogen) | diverse serotypes |
| 937 | 1           | 0.36 | +           | - (ExPEC, non pathogen) | O43:H2 |
| 1385| 1           | 0.36 | -           | - (APEC)         | Ounk:H4         |
| 1611| 2           | 0.72 | -           | - (APEC)         | diverse serotypes |
| 5082| 1           | 0.36 | -           | NR               | NR              |
| 332 | 1           | 0.36 | -           | NR               | O171:H2         |
| 5395| 1           | 0.36 | -           | NR               | O74:H8          |
| 5435| 2           | 0.72 | -           | NR               | Ounk:H16        |
| 5530| 1           | 0.36 | -           | NR               | Ounk:H21        |
| 5602| 2           | 0.72 | -           | NR               | 36:H28          |
| 5960| 1           | 0.36 | -           | NR               | NR              |
| 5975| 1           | 0.36 | -           | NR               | O113:H21        |
| 6509| 1           | 0.36 | -           | NR               | O168:H8         |
| 6632| 1           | 0.36 | -           | NR               | O8:H16          |
| 6638| 1           | 0.36 | -           | NR               | Ounk:H19        |
| 3017| 1           | 0.36 | -           | NR               | O116:H21        |
| 6640| 1           | 0.36 | -           | NR               | O113:H21        |
| 6641| 1           | 0.36 | -           | NR               | O130:H11        |
| 6642| 1           | 0.36 | -           | NR               | O113:H21        |
| 1112| 1           | 0.36 | -           | - (nonpathogen)  | diverse serotypes |
| 1176| 1           | 0.36 | -           | - (nonpathogen)  | O36:H14         |
| 2161| 1           | 0.36 | -           | - (nonpathogen)  | O180:H14        |
| 2217| 4           | 1.44 | -           | - (nonpathogen)  | diverse serotypes |
| 2389| 1           | 0.36 | -           | - (nonpathogen)  | NR              |
| 1086| 10          | 3.60 | -           | - (nonpathogen)  | diverse serotypes |
| 2385| 7           | 2.52 | -           | - (nonpathogen)  | O8:H19          |

(Continued)
spanning tree allows visualization of allele differences between strains with the same ST that was not seen with the NJ tree (Fig 2).

eae positive Non-STEC strains virulence gene profiles
Among the 55 non-STECs (lacking either stx gene by in silico analysis) strains isolated from FDA regulated foods, we found 35 that were positive for the eae gene (S4 Table). The majority were classified as atypical EPEC (aEPEC) eae\textsuperscript{+} and bfpA\textsuperscript{−}. Two of them (IEH-NGS-ECO-00094, and IEH-NGS-ECO-00100) carried bfpA\textsuperscript{−} -the first gene of the bfp operon that encodes a type IV pilus- (eae\textsuperscript{+} and bfpA\textsuperscript{−}) but were missing most of the common genes found in typical EPEC (S4 Table, typical EPEC lineage 1 strain E2348/69). Therefore, we classified them as aEPEC.

Discussion
STECs are the most dangerous among diarrheagenic E. coli affecting public health worldwide [5,7,23,27,42,43]. Usually the most threatening STEC are those of O157:H7 serotype [44,45]. However, in recent years there has been an increase in the occurrence of many non-O157 serotypes in humans associated with consumption of contaminated food, including produce and other FDA regulated products [2,46,47]. Some studies have characterized STECs presence and their virulence potential from FDA regulated products [20,23,48]. Most of the STEC isolated from those products have been only initially screened for the presence of some virulence genes using PCR [23,49]. In the present study, we performed an in-depth analysis with whole genome sequencing of 331 presumptive STEC strains. These strains were isolated from FDA regulated foods recovered during a period of 2003–2017 by two surveillance programs (FDA ORA, and MDP USDA). STECs were isolated from 22 food commodities. It is worth mentioning that even though the sampling did not occur in all states, the food commodities had nationwide (or at least multistate) distribution.

The STEC analyzed in this study were isolated from a wide variety of foods (Table 1), with the majority isolated from spinach (32%), flour (21%), lettuce (13%), and cilantro (12%) samples during the period 2010–2017. The actual frequency of flour STECs should be assessed at a lower frequency of 9%; the spike observed in their frequency was due to the outbreak in flour in 2016, where most STECs (37 strains (66%) of total flour STECs) were isolated. A better reflection of the frequency of STEC isolated per food commodity, specifically produce, can be found in Feng and Reddy (2013) [23]. Nevertheless, the presence of STECs in FRFDA per year remained relatively low, with a median of 30 isolates per year. As pointed previously, these variations in frequency of isolation can be due to seasonal and geographical variations, or to sporadic outbreaks as was observed for O121:H19 STEC strains isolated from flour in 2016 (23).
Table 4. *in silico* characterization of STECs from this study for presence of virulence genes and their serotype.

| strains          | ST | serotype | stx1 type | stx2 type | eae type | ehxA | espP | etpD | toxB | katP | subA | saa | sab |
|------------------|----|----------|-----------|-----------|----------|------|------|------|------|------|------|-----|-----|
| CFSAN041120      | 10 | O2:H27   | -         | a         | -        | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051538      | 10 | Ounk:H32 | a         | -         | -        | -    | -    | +    | -    | -    | -    | -   | -   |
| CFSAN046715      | 11 | O157:H7  | -         | a         | gamma-1  | +    | +    | +    | -    | -    | -    | -   | -   |
| CFSAN046720      | 11 | O157:H7  | -         | a         | gamma-1  | +    | +    | +    | -    | -    | -    | -   | -   |
| FDA00009839      | 11 | O157:H7  | -         | a         | gamma-1  | +    | +    | +    | -    | -    | -    | -   | -   |
| CFSAN053336      | 16 | O111:H8  | a         | -         | theta-2  | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN053346      | 16 | O111:H8  | a         | a         | theta-2  | +    | -    | -    | -    | -    | -    | -   | -   |
| FDA00010429      | 17 | O103:H2  | a         | -         | epsilon  | +    | -    | -    | -    | -    | -    | -   | -   |
| IEH-NGS-ECO-00075| 17 | O103:H2  | a         | -         | epsilon  | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046724      | 21 | O26:H11  | a         | -         | beta-1   | +    | +    | -    | -    | -    | -    | -   | -   |
| CFSAN046724      | 21 | O26:H11  | a         | -         | beta-1   | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN053342      | 21 | O103:H11 | a         | -         | beta-1   | +    | +    | -    | +    | +    | -    | -   | -   |
| CFSAN053343      | 21 | O26:H11  | a         | -         | beta-1   | +    | -    | -    | +    | +    | -    | -   | -   |
| CFSAN053345      | 21 | O26:H11  | a         | -         | beta-1   | +    | -    | -    | +    | +    | -    | -   | -   |
| FDA00010430      | 21 | O26:H11  | a         | -         | beta-1   | +    | -    | -    | +    | +    | -    | -   | -   |
| IEH-NGS-ECO-00076| 21 | O26:H11  | a         | -         | beta-1   | +    | -    | -    | +    | +    | -    | -   | -   |
| IEH-NGS-ECO-00213| 21 | O26:H11  | a         | -         | beta-1   | +    | -    | -    | +    | +    | -    | -   | -   |
| IEH-NGS-ECO-00227| 25 | O128ac:H2| c         | -         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| IEH-NGS-ECO-00225| 29 | Ounk:H11 | -         | a         | beta-1   | +    | -    | -    | -    | -    | -    | -   | -   |
| IEH-NGS-ECO-00224| 32 | O145:H28 | a         | d         | gamma-1  | +    | +    | -    | +    | +    | -    | -   | -   |
| CFSAN051773      | 33 | O91:H14  | a         | d         | -        | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN053344      | 33 | O91:H14  | a         | d         | -        | +    | -    | -    | -    | -    | +    | +   | +   |
| IEH-NGS-ECO-00232| 40 | Ounk:H21 | c         | -         | -        | -    | -    | -    | +    | -    | -    | -   | -   |
| CFSAN051526      | 43 | O6:H10   | c         | -         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051527      | 43 | O6:H10   | c         | -         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051533      | 43 | O6:H10   | c         | -         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051535      | 43 | O6:H10   | c         | -         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051537      | 43 | O6:H10   | c         | -         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046659      | 58 | O116:H21 | -         | a         | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN046700      | 88 | O8:H9    | -         | e         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051531      | 88 | O8:H30   | -         | e         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046737      | 101| O82:H8   | -         | a         | -        | +    | +    | -    | -    | -    | -    | +   | +   |
| CFSAN046738      | 101| O82:H8   | -         | a         | -        | +    | +    | -    | -    | -    | -    | +   | +   |
| CFSAN046749      | 101| O21:H21  | -         | a         | -        | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046750      | 119| O165:H28 | a         | a         | epsilon-2| +    | +    | -    | +    | +    | -    | -   | -   |
| CFSAN046672      | 154| O88:H25  | a         | a         | -        | +    | -    | -    | -    | -    | -    | +   | +   |
| CFSAN046673      | 154| O88:H25  | a         | a         | -        | +    | -    | -    | -    | -    | -    | +   | +   |
| CFSAN046682      | 154| O134:H38 | a         | d         | -        | +    | -    | -    | -    | -    | -    | +   | +   |
| CFSAN056112      | 154| O88:H25  | a         | a         | -        | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN056113      | 154| O88:H25  | a         | a         | -        | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN051557      | 156| O174:H28 | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011520      | 162| O8:H19   | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046788      | 173| O181:H49 | a         | a         | -        | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046740      | 173| O181:H49 | -         | d         | -        | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046747      | 173| O181:H49 | -         | d         | -        | +    | -    | -    | -    | -    | +    | +   | +   |
| IEH-NGS-ECO-00108| 173| O181:H49 | -         | a         | -        | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046751      | 173| O181:H49 | -         | d         | -        | +    | -    | -    | -    | -    | +    | +   | +   |

(Continued)
| Strains          | ST     | Serotype          | Stx1 Type | Stx2 Type | Eae Type | EhxA | EspP | EtpD | ToxB | KatP | SubA | Saa | Sab |
|------------------|--------|-------------------|-----------|-----------|-----------|------|------|------|------|------|------|-----|-----|
| CFSAN041109      | 205    | Ounk:H19          | -         | -         | +         | -    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046650      | 205    | O153/O178:H19     | -         | -         | +         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046690      | 205    | O153/O178:H19     | a         | a         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046704      | 205    | Ounk:H19          | -         | a         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN051550      | 205    | Ounk:H19          | -         | a         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN051551      | 205    | Ounk:H19          | -         | a         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| FDA00009731      | 205    | O153/O178:H19     | a         | d         | -         | +    | -    | -    | -    | -    | -    | +   | +   |
| CFSAN046660      | 223    | O113:H21          | -         | a         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN046665      | 223    | O113:H21          | -         | a         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN053329      | 223    | O113:H21          | -         | a         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN046734      | 223    | O113:H21          | -         | a         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN041114      | 295    | Ounk:H16          | -         | a         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046640      | 295    | Ounk:H16          | c         | b         | -         | -    | -    | -    | -    | -    | -    | +   | +   |
| CFSAN046691      | 295    | Ounk:H11          | a         | -         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN053347      | 295    | Ounk:H16          | -         | a         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN046735      | 295    | Ounk:H11          | -         | a         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN041108      | 297    | O130:H11          | a         | d         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046639      | 297    | O130:H11          | a         | -         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046719      | 297    | O130:H11          | -         | d         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046753      | 297    | O179:H8           | a         | -         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN051516      | 297    | O130:H11          | a         | -         | -         | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN051524      | 297    | O130:H11          | a         | d         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN051549      | 297    | O130:H11          | a         | d         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN056111      | 297    | O130:H11          | -         | a         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| IEH-NGS-ECO-00087| 297    | O179:H8           | a         | -         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046717      | 306    | O98:H21           | a         | -         | zeta      | +    | +    | -    | -    | -    | -    | -   | -   |
| CFSAN051525      | 306    | O98:H21           | a         | -         | zeta      | +    | +    | -    | -    | -    | -    | -   | -   |
| CFSAN051544      | 325    | O15:H16           | g         | -         | -         | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046643      | 329    | O136:H16          | a         | -         | -         | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN053337      | 329    | O136:H16          | a         | -         | -         | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046636      | 332    | O171:H2           | c         | -         | -         | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN053335      | 342    | O5:Hunk           | a         | -         | beta-1    | +    | +    | -    | -    | -    | -    | -   | -   |
| IEH-NGS-ECO-00230| 342    | O5:Hunk           | a         | -         | beta-1    | +    | +    | -    | -    | -    | -    | -   | -   |
| CFSAN053330      | 394    | O17/O77:H18       | d         | -         | +         | +    | -    | -    | -    | -    | +    | +   | +   |
| IEH-NGS-ECO-00231| 398    | O136:H20          | c         | -         | -         | -    | -    | -    | -    | -    | +    | -   | -   |
| CFSAN046746      | 442    | O91:H21           | a         | -         | -         | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051529      | 442    | O146:H21          | c         | b         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN051540      | 442    | O146:H21          | c         | -         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046688      | 443    | O153/O178:H19     | a         | d         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046752      | 443    | O153/O178:H19     | a         | d         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046631      | 446    | O22:H8            | -         | c         | -         | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00009425      | 446    | O22:H8            | -         | c         | -         | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN053334      | 515    | Ounk:H29          | -         | d         | -         | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046632      | 641    | O117:H10          | a         | -         | -         | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046646      | 642    | O187:H52          | c         | -         | -         | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046648      | 642    | O187:H52          | c         | -         | -         | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046649      | 642    | O187:H52          | c         | -         | -         | -    | -    | -    | -    | -    | -    | -   | -   |

(Continued)
| Strains           | ST     | Serotype | stx1 Type | stx2 Type | eae Type | ehxA | espP | etpD | toxB | katP | subA | saa | sab |
|-------------------|--------|----------|-----------|-----------|----------|------|------|------|------|------|------|-----|-----|
| CFSAN046651       | 655    | O121:H19 |           |           |          |      |      |      |      |      |      |     |     |
| FDA00010253       | 655    | O121:H19 | -         | a         | epsilon-2 | +    | -    | -    | -    | -    | -    | -   | -   |
| FDA00010254       | 655    | O121:H19 | -         | a         | epsilon-2 | +    | +    | -    | -    | -    | -    | -   | -   |
| FDA00010255       | 655    | O121:H19 | -         | a         | epsilon-2 | +    | +    | -    | -    | -    | -    | -   | -   |
| FDA00010256       | 655    | O121:H19 | -         | a         | epsilon-2 | +    | +    | -    | -    | -    | -    | -   | -   |
| FDA00010257       | 655    | O121:H19 | -         | a         | epsilon-2 | +    | +    | -    | -    | -    | -    | -   | -   |
| FDA00010258       | 655    | O121:H19 | -         | a         | epsilon-2 | +    | +    | -    | -    | -    | -    | -   | -   |
| FDA00010259       | 655    | O121:H19 | -         | a         | epsilon-2 | +    | +    | -    | -    | -    | -    | -   | -   |
| FDA00010260       | 655    | O121:H19 | -         | a         | epsilon-2 | +    | +    | -    | -    | -    | -    | -   | -   |
| IEH-NGS-E CO-00103| 678    | O183:H18 | a         | d         |          |      |      |      |      |      |      |     |     |
| CFSAN041119       | 661    | O174:H2  | -         | c         | -        | +    | +    | -    | -    | -    | -    | -   | +   |
| CFSAN051547       | 661    | O174:H2  | -         | c         | -        | +    | +    | -    | -    | -    | -    | -   | +   |
| FDA00011331       | 662    | O17/07:H45| a         | d         |          |      |      |      |      |      |      | +   | +   |
| FDA00011332       | 662    | O17/07:H45| a         | d         |          |      |      |      |      |      |      | +   | +   |
| CFSAN051530       | 675    | O76:H19  | c         | -         |          | +    | +    | -    | -    | -    | -    | +   | +   |
| CFSAN046652       | 677    | Ounk:H21 | -         | d         |          |      |      |      |      |      |      | -   | -   |
| CFSAN04748        | 677    | O174:H21 | a         | d         |          | +    | +    | -    | -    | -    | -    | +   | +   |
| CFSAN051519       | 677    | O174:H21 | -         | a         |          | -    | -    | -    | -    | -    | -    | -   | +   |

(Continued)
| strains                  | ST      | serotype | stx1 type | stx2 type | eae type | ehxA | espP | etpD | toxB | katP | subA | saa | sab |
|-------------------------|---------|----------|-----------|-----------|----------|------|------|------|------|------|------|-----|-----|
| CFSAN051520             | 677     | O174:H21 | -         | a         | -        | -    | -    | -    | -    | -    | -    | -   | +   |
| IEH-NGS-ECO-00080       | 677     | O174:H21 | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046699             | 679     | O163:H19 | -         | d         | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN046705             | 679     | O163:H19 | -         | a         | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN051541             | 679     | O163:H19 | a         | d         | -        | +    | +    | -    | -    | +    | +    | +   | +   |
| CFSAN053338             | 691     | Ounk:H20 | a         | d         | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN046709             | 691     | Ounk:H20 | a         | -         | -        | +    | +    | -    | -    | -    | -    | +   | +   |
| CFSAN041113             | 692     | O74:H42  | a         | d         | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN051548             | 692     | O74:H42  | a         | d         | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN051553             | 692     | O74:H42  | a         | d         | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN046689             | 706     | O32:H1   | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046666             | 718     | O168:H8  | -         | a         | -        | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046693             | 718     | O168:H8  | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046725             | 718     | O168:H8  | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00010428             | 723     | O103:H11 | a         | -         | beta-1   | -    | -    | -    | +    | -    | -    | -   | -   |
| FDA00010431             | 723     | O103:H11 | a         | -         | beta-1   | +    | -    | -    | -    | +    | -    | -   | -   |
| FDA00010457             | 723     | O103:H11 | a         | -         | beta-1   | +    | +    | -    | -    | -    | -    | -   | -   |
| IEH-NGS-ECO-00223       | 723     | O103:H11 | a         | -         | beta-1   | +    | +    | -    | +    | +    | -    | -   | -   |
| CFSAN046707             | 724     | Ounk:H20 | a         | -         | -        | +    | +    | -    | -    | +    | +    | +   | +   |
| CFSAN046733             | 747     | O17/O77:H45 | -       | a        | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN046741             | 747     | O17/O77:H45 | -       | a        | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN046742             | 747     | O17/O77:H45 | -       | a        | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN046743             | 747     | O17/O77:H45 | -       | a        | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| FDA00010882             | 747     | O17/O77:H45 | -       | a        | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| FDA00010883             | 747     | O17/O77:H45 | -       | a        | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| FDA00010884             | 747     | O17/O77:H45 | -       | a        | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| FDA00010885             | 747     | O17/O77:H45 | -       | a        | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| FDA000104116            | 906     | O74:H8   | -         | d         | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN045555             | 906     | O74:H8   | -         | d         | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| IEH-NGS-ECO-00096       | 906     | O74:H8   | -         | a         | -        | +    | +    | -    | -    | -    | +    | +   | +   |
| CFSAN051545             | 937     | O43:H2   | a         | a         | -        | +    | -    | -    | -    | -    | -    | +   | +   |
| CFSAN046721             | 942     | O116:H28 | a         | -         | +        | +    | -    | -    | -    | -    | +    | +   | +   |
| CFSAN046713             | 955     | O139:H1  | -         | e         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN045559             | 993     | O106:H30 | -         | e         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046730             | 993     | O106:H30 | -         | e         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011815             | 1086    | O8:H14   | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011816             | 1086    | O8:H14   | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011817             | 1086    | O8:H14   | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011818             | 1086    | O8:H14   | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011819             | 1086    | O8:H14   | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011820             | 1086    | O8:H14   | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011821             | 1086    | O8:H14   | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011822             | 1086    | O8:H14   | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011823             | 1086    | O8:H14   | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011824             | 1086    | O8:H14   | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00011218             | 1112    | O142:H27 | -         | e         | -        | -    | -    | -    | -    | -    | -    | -   | -   |

(Continued)
| Strains   | ST      | Serotype | STX1 Type | STX2 Type | Eae Type | EthA | EspP | EtpD | ToxB | KatP | SubA | Saa | Sab |
|-----------|---------|----------|-----------|-----------|----------|------|------|------|------|------|------|-----|-----|
| CFSAN046653 | 1176    | O36:H14  | -         | g         | -        | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046655 | 1385    | Unknown  | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046683 | 1431    | O8:H19   | -         | d/e       | -        | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046728 | 1611    | O159:H19 | -         | d/e       | -        | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051515 | 1611    | O159:H19 | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046668 | 1727    | Unknown  | -         | c         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046669 | 1727    | Unknown  | -         | c         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046671 | 1727    | Unknown  | -         | c         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046692 | 1727    | Unknown  | -         | c         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| FDA00010    | 1792    | O111:H8  | a         | a         | -        | theta-2 | +    | -    | -    | -    | -    | -    | -   | -   |
| IEH-NGS-ECO-00220 | 1817 | O104:H7  | c         | -         | -        | -    | -    | -    | -    | +    | -    | -   | -   |
| IEH-NGS-ECO-00221 | 1967 | O103:H2  | a         | -         | epsilon  | +    | -    | -    | -    | +    | -    | -   | -   |
| IEH-NGS-ECO-00221 | 1967 | O103:H2  | a         | -         | epsilon  | +    | -    | -    | -    | +    | -    | -   | -   |
| CFSAN041110 | 2008    | Unknown  | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN041111 | 2008    | Unknown  | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN041112 | 2008    | Unknown  | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046642 | 2008    | Unknown  | -         | a         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046657 | 2008    | Unknown  | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046684 | 2008    | Unknown  | -         | d/c       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046685 | 2008    | Unknown  | -         | d/c       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046686 | 2008    | Unknown  | -         | d/c       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046687 | 2008    | Unknown  | -         | d/c       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046708 | 2008    | Unknown  | -         | d/c       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046710 | 2008    | Unknown  | -         | d/c       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051552 | 2008    | Unknown  | -         | d/c       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046701 | 2161    | O180:H14 | -         | d/e       | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046641 | 2217    | O45:H16  | a         | -         | -        | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046712 | 2217    | O76:H21  | a         | -         | -        | -    | -    | -    | -    | -    | -    | -   | +   |
| CFSAN046714 | 2217    | O8:H16   | a         | -         | -        | +    | -    | -    | -    | -    | -    | -   | +   |
| CFSAN051522 | 2217    | O84:H38  | a         | -         | -        | +    | -    | -    | -    | -    | -    | -   | +   |
| CFSAN046723 | 2385    | O8:H19   | a         | a         | -        | +    | -    | -    | -    | +    | +    | -   | +   |
| CFSAN051517 | 2385    | O8:H19   | a         | a         | -        | +    | -    | -    | -    | +    | +    | -   | +   |
| CFSAN051518 | 2385    | O8:H19   | a         | a         | -        | +    | -    | -    | -    | +    | +    | -   | +   |
| IEH-NGS-ECO-00082 | 2385 | O8:H19   | a         | a         | -        | +    | -    | -    | -    | +    | +    | -   | +   |
| IEH-NGS-ECO-00088 | 2385 | O8:H19   | a         | a         | -        | +    | -    | -    | -    | +    | +    | -   | +   |
| IEH-NGS-ECO-00089 | 2385 | O8:H19   | a         | a         | -        | +    | -    | -    | -    | +    | +    | -   | +   |
| IEH-NGS-ECO-00090 | 2385 | O8:H19   | a         | a         | -        | +    | -    | -    | -    | +    | +    | -   | +   |
| CFSAN041115 | 2387    | O185:H7  | -         | c         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN041118 | 2387    | O185:H7  | -         | c         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046670 | 2387    | O185:H7  | -         | c         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051528 | 2387    | O185:H7  | -         | c         | -        | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051546 | 2387    | O185:H7  | -         | c         | -        | +    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN051554 | 2387    | O185:H7  | -         | c         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046744 | 2387    | O185:H7  | -         | c         | -        | -    | -    | -    | -    | +    | +    | -   | +   |
| CFSAN046745 | 2387    | O185:H7  | -         | c         | -        | +    | -    | -    | -    | +    | +    | -   | +   |
| CFSAN053339 | 2388    | O15:H27  | -         | d         | -        | -    | -    | -    | -    | -    | -    | -   | -   |
| CFSAN046656 | 2389    | Unknown  | -         | d         | -        | +    | -    | -    | -    | +    | +    | +   | +   |

(Continued)
WGS revealed that these STECs were highly variable with the existence of 95 different sequence types (STs) and belonging to at least 81 different serotypes. Some serotypes could not be predicted and might be due to the fact that the O and H type genes were not present in the database used which includes the most frequent serotypes found in clinical cases. Most STs were observed only once while some others were observed more frequently. ST655 was

Table 4. (Continued)

| strains         | ST  | serotype    | stx1 type | stx2 type | eae type | ehxA | espP | etpD | toxB  | katP | subA | saa | sab |
|-----------------|-----|-------------|-----------|-----------|----------|------|------|------|-------|------|------|-----|-----|
| CFSAN046679     | 2520| O116:H49    | -         | a         | -        | +    | +    | -    | -     | -    | +    | +   | +   |
| CFSAN046731     | 2520| O116:H49    | -         | a         | -        | +    | +    | -    | -     | +    | +    | +   | +   |
| IEH-NGS-ECO-00204 | 2520| O116:H49    | a         | a         | -        | +    | +    | -    | -     | +    | +    | +   | +   |
| CFSAN046633     | 3017| O116:H21    | a         | a         | -        | +    | -    | -    | -     | +    | +    | +   | +   |
| IEH-NGS-ECO-00070 | 3759| O8:H49      | -         | a         | -        | +    | -    | -    | -     | +    | +    | +   | +   |
| CFSAN046667     | 4173| O79:H2      | a         | -         | -        | -    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046676     | 4173| O79:H2      | a         | -         | -        | -    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046637     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046644     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046654     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046661     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046680     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046695     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046718     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046726     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046727     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| FDA00009866     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| FDA00009867     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046729     | 4496| O8:H28      | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| FDA00011519     | 5082| O180:H14    | -         | d/e       | -        | -    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046674     | 5299| O8:H49      | a         | -         | -        | -    | -    | -    | -     | -    | -    | +   | +   |
| CFSAN046675     | 5299| O8:H49      | a         | -         | -        | +    | +    | -    | -     | -    | -    | +   | +   |
| CFSAN046722     | 5395| O74:H8      | a         | -         | -        | +    | +    | -    | -     | -    | +    | +   | +   |
| CFSAN046694     | 5435| Ounk:H16    | -         | a         | -        | +    | -    | -    | -     | +    | +    | +   | +   |
| CFSAN046696     | 5435| Ounk:H16    | -         | a         | -        | +    | +    | -    | -     | +    | +    | +   | +   |
| CFSAN046664     | 5530| Ounk:H21    | -         | d/e       | -        | -    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN041117     | 5602| O36:H28     | -         | g         | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN051556     | 5602| O36:H28     | -         | g         | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| IEH-NGS-ECO-00105 | 5960| Ounk:H19    | -         | a         | -        | +    | -    | -    | -     | +    | +    | +   | +   |
| CFSAN046702     | 5973| Ounk:H2     | -         | d         | -        | +    | -    | -    | -     | +    | +    | +   | +   |
| CFSAN046716     | 5973| Ounk:H2     | -         | c         | -        | +    | -    | -    | -     | +    | +    | +   | +   |
| CFSAN046739     | 5975| O113:H21    | -         | a         | -        | +    | +    | -    | -     | +    | +    | +   | +   |
| CFSAN046645     | 6475| O17/O77:H45 | a         | a         | -        | +    | -    | -    | -     | -    | +    | +   | +   |
| CFSAN046647     | 6475| O17/O77:H45 | a         | a         | -        | +    | -    | -    | -     | -    | +    | +   | +   |
| CFSAN046663     | 6509| O168:H8     | -         | d/e       | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN051521     | 6632| O8:H16      | -         | d         | -        | +    | -    | -    | -     | -    | +    | +   | +   |
| CFSAN046662     | 6638| Ounk:H19    | -         | a         | -        | +    | -    | -    | -     | +    | +    | +   | +   |
| CFSAN046711     | 6639| O174:H21    | -         | c         | -        | +    | -    | -    | -     | -    | -    | -   | -   |
| CFSAN046677     | 6640| O113:H21    | -         | d         | -        | +    | +    | -    | -     | +    | +    | +   | +   |
| CFSAN046706     | 6641| O130:H11    | a         | -         | -        | +    | -    | -    | -     | -    | +    | +   | +   |
| CFSAN046697     | 6642| O113:H21    | a         | -         | +        | -    | -    | -    | -     | +    | +    | +   | +   |

https://doi.org/10.1371/journal.pone.0214620.t004
observed up to 38 times among the STECs analyzed and it was because 37 of those STEC strains were recovered during the flour outbreak in 2016 [20]. According to what was found in

| Strains                  | aadA | aph3 | strA | strB | bla TEM | mef(B) | floR | QnrB | sul1 | sul2 | sul3 | tetA | tetB | tetC | dfrA |
|--------------------------|------|------|------|------|---------|--------|------|------|------|------|------|------|------|------|------|
| CFSAN053338              | -    | +    | +    | -    | -       | -      | -    | -    | -    | -    | -    | -    | -    | -    | -    |
| CFSAN046714              | -    | -    | -    | -    | -       | +      | +    | -    | -    | -    | +    | -    | +    | -    | -    |
| CFSAN051552              | -    | +    | +    | -    | -       | -      | +    | -    | +    | -    | -    | +    | -    | -    | -    |
| CFSAN046710              | -    | -    | +    | -    | -       | -      | -    | -    | -    | -    | +    | -    | -    | -    | -    |
| CFSAN046642              | -    | -    | +    | -    | -       | -      | +    | -    | +    | -    | -    | +    | -    | -    | -    |
| IEH-NGS-ECO-00231       | -    | -    | -    | -    | -       | -      | -    | -    | -    | +    | -    | -    | -    | -    | -    |
| CFSAN053336              | -    | +    | +    | -    | -       | -      | -    | -    | -    | +    | -    | -    | -    | -    | -    |
| FDA00009425              | -    | -    | +    | -    | -       | +      | -    | -    | -    | +    | -    | -    | -    | -    | -    |
| FDA00011218              | -    | -    | -    | -    | -       | -      | -    | -    | -    | -    | -    | -    | -    | +    | -    |
| IEH-NGS-ECO-             | +    | -    | +    | -    | -       | -      | -    | -    | +    | +    | -    | -    | -    | -    | -    |
| CFSAN051526              | +    | -    | -    | -    | -       | -      | +    | -    | +    | -    | -    | +    | -    | -    | -    |
| CFSAN046636              | -    | -    | +    | -    | -       | -      | +    | -    | -    | -    | +    | -    | -    | -    | -    |
| CFSAN046669              | -    | +    | +    | -    | -       | -      | -    | -    | -    | +    | -    | -    | -    | -    | -    |
| CFSAN046668              | -    | -    | +    | -    | -       | -      | +    | -    | +    | -    | -    | -    | -    | -    | -    |
| CFSAN053334              | -    | -    | +    | -    | +       | -      | +    | -    | +    | -    | -    | -    | -    | -    | -    |
| CFSAN046687              | -    | -    | +    | -    | -       | -      | +    | -    | +    | -    | -    | -    | -    | -    | -    |
| CFSAN051521              | +    | -    | -    | -    | -       | -      | -    | -    | -    | -    | +    | -    | -    | -    | -    |
| CFSAN046730              | -    | +    | +    | -    | -       | -      | +    | -    | -    | +    | -    | +    | -    | -    | -    |
| CFSAN041112              | -    | +    | +    | -    | -       | -      | +    | -    | -    | +    | -    | +    | -    | -    | -    |
| CFSAN041111              | +    | -    | +    | -    | -       | -      | +    | -    | -    | +    | -    | +    | -    | -    | -    |
| CFSAN046671              | -    | -    | +    | -    | -       | -      | +    | -    | -    | +    | -    | +    | -    | -    | -    |
| CFSAN046708              | -    | -    | +    | -    | -       | -      | +    | -    | -    | +    | -    | +    | -    | -    | -    |
| CFSAN046685              | -    | -    | +    | -    | -       | -      | -    | -    | +    | -    | -    | +    | -    | -    | -    |
| CFSAN046684              | -    | -    | +    | -    | -       | -      | -    | -    | +    | -    | -    | +    | -    | -    | -    |
| CFSAN051527              | +    | -    | -    | -    | -       | -      | +    | -    | -    | +    | -    | +    | -    | -    | -    |
| CFSAN046686              | -    | -    | +    | -    | -       | -      | -    | -    | +    | -    | -    | +    | -    | -    | -    |
| CFSAN051535              | -    | -    | +    | +    | -       | -      | -    | -    | -    | +    | -    | +    | -    | -    | -    |
| CFSAN046693              | -    | -    | +    | -    | -       | -      | -    | -    | +    | -    | -    | +    | -    | -    | -    |
| CFSAN046713              | -    | -    | +    | +    | -       | -      | +    | -    | -    | +    | -    | +    | -    | -    | -    |
| CFSAN046725              | -    | -    | -    | +    | -       | -      | -    | -    | -    | +    | -    | -    | -    | -    | -    |
| CFSAN051531              | -    | -    | -    | +    | -       | -      | -    | -    | -    | +    | -    | -    | -    | -    | -    |
| CFSAN051539              | +    | +    | -    | -    | +       | +      | -    | -    | -    | -    | -    | -    | -    | -    | -    |
| CFSAN041110              | -    | -    | +    | -    | -       | -      | -    | -    | +    | -    | -    | +    | -    | -    | -    |

*a* Aminoglycoside,  
*b* Beta-lactamase,  
*c* Macrolide,  
*d* Phenicol,  
*e* Quinolone,  
*f* Sulphonamide,  
*g* Tetracycline, and  
*h* Trimethoprim.  

*strain carrying blaOXA*\(^b\)* gene.  
*strain carrying blaCMY-2*\(^b\)* gene.  
*strain carrying cml*\(^d\) and *cmlA1*\(^d\) genes.

https://doi.org/10.1371/journal.pone.0214620.t005
Enterobase (http://enterobase.warwick.ac.uk), the majority of the STEC STs observed in this study [69/95–73%] had been reported as causing disease in humans. Furthermore, of these potential human pathogenic STECs, strains belonging to 18 of those STs (19%) were additionally associated with strains causing EHEC-related illnesses (Table 3). Among the known ST associated with causing HC illnesses or HUS cases we found: ST21 and ST29 (O26:H11), ST11 (O157:H7), ST33 (O91:H14), ST17 (O103:H2), and ST16 (O111:H-), among others [7,42,50,51]. There are some STs, that have been described as EAEC/ETEC, however in our dataset, the strains with those STs did not carry any virulence genes for those pathotypes. These results emphasized the need to analyze the data as a whole (e.g. to include virulence genes); ST alone is not an accurate predictor of pathotype or virulence.

Although some strains have the same ST, they show differences in their virulence profile as well as their Shiga toxin gene content. For example, there were 5 strains that were ST10 and from these only 2 were classified as STECs, with one carrying stx1a while the other carried stx2a. Both were negative for any of the attaching genes (eae, or saa genes), and therefore considered low risk for causing infection in a healthy individual. This demonstrates that a single characteristic (e.g. ST or serotype) is not enough to make an inference of the potential pathogenic trait of any STECs (http://www.fao.org/documents/card/en/c/CA0032EN). The better
way is to take all the information into consideration (ST, stx type, attaching genes, serotype, etc) in order to make a more informed prediction of the pathogenic potential of any STEC in conjunction with historical available data on clinical cases. For example, a strain of O113:H21 stx2a positive that doesn’t possess eae but has saa (a gene that encodes an auto-agglutinating adhesion) and has been linked to HUS cases [14], could be potentially harmful to humans. A similar analysis could be done in the case of any STEC that has all those attributes but has not been linked to any human cases. Even though we cannot predict the actual outcome of an infection with this strain, it still warrants a warning about its presence in foods that are consumed raw as is the case with fresh produce.

We tested for 95 known virulence genes [27] found in the most common E. coli pathotypes and did not find any genes present that would characterize the strains as STEC/EAEC//ETEC/EIEC hybrids. Among the adherence factors, eae and saa genes were found in 24% and 26% of the STEC strains, respectively. Strains that carry eae did not carry saa, and vice versa, as previously observed for STEC isolated from fresh produce [23]. Regarding the presence of Shiga toxin type, there was great variation with most strains (67%) carrying only stx type 2, 19% carrying only stx type 1 while 15% carried both stx types. Among the stx2 there were 144 that were either a,d, or c variants, which are the stx2 variants found among clinical cases [31,32,52–54] and that have a specific tropism for humans [53]. The remaining 40 STECs carrying stx type 2 alone were stx variants e, d/e and g which have been found in animal reservoirs [55]. The remaining putative virulence genes were sporadically found with the most common exhA gene found in 61% of the STECs, while espP was found in 43% of the STECs. These two genes can be found in the virulence plasmid and appear to participate in STECs infection in humans [9,11,15]. In summary, 46 of the STECs analyzed in this study carried a combination of stx2a/d and eae genes. Many researchers and food safety institutions considered a STEC with this virulence gene combination to possess an elevated risk to humans [22,53]. On the other hand, 94 of these STECs carried a combination of stx2a/d/c and saa genes. In this case, the risk is harder to determine and are considered of moderate risk if serotype was previously found associated with human cases.
We also confirmed the presence of antimicrobial resistance (AMR) genes in some of the analyzed STEC strains with a low prevalence (12%). However, those few strains carried multiple antimicrobial resistance genes. The presence of strains carrying multiple AMR genes is worrisome since they can be shared amongst other *E. coli* and could possible participate in the dissemination of AMR in their environments, as has been observed for tetracycline genes in *E. coli* isolates from beef cattle [56], for colistin resistance (*mcr-1* gene) through plasmid-mediated transfer [57], and for ampicillin resistance genes in *E. coli* in an infant treated with antibiotics [58].

Phylogenetic analysis by a custom cgMLST analysis of these 276 STECs confirmed the MLST *in silico* analysis, with many different defined clades among these STECs isolated from FRFDA. The cgMLST analysis is a fast method of analysis and provides an initial visualization of the relationships among the strains analyzed. Comparable results have been observed for establishing fast relationships among genomes from diverse bacterial pathogens [27,29,59–64]. A further analysis using only the genomes of strains that are located within each individual or among selected clades can produce a more detailed evolutionary history, using single nucleotide analyses. This SNP analysis can help in determining the potential source, phylogenetic nature, lineage, and timeline of transmission of each group, as has been shown for the ST36 lineage of *Vibrio parahaemolyticus* [65].

EPECs are the leading cause of infantile diarrhea in developing countries [66,67]. EPEC strains do not carry *stx* genes but typical EPECs (tEPEC) have *eae* and *bfp* genes, and their main reservoir is humans [68]. The *eae* gene is located in the chromosome, in the LEE operon, while the *bfp* operon is typically located in the large EPEC adherence factor (EAF) virulence plasmid [68]. These tEPEC also carry the *perA* gene, which increases the expression of LEE elements [68,69]. Interestingly, 11% of our presumptive STECs were shown by *in silico* analysis to be atypical EPECs (aEPEC). Among their unusual features are the absence of the EAF plasmid, and their reservoirs can be animals or humans [68]. It is possible these aEPECs might had lost their phages upon culturing, as this pattern has been observed in clinical isolates of *E. coli* upon sub-cultivation [70]. The aEPEC we observed in this study may have the capacity to produce A/E lesions, since they carried both the *eae* and *tir* gene, which are the effector and receptor necessary for the formation of the A/E lesion [71].

Our results suggest that finding aEPECs in food could be of particular concern, as these strains have the potential for acquiring the *stx* phage, as observed in the *E. coli* O104:H4 strain found in Germany [72]. That strain was an entero-aggregative *E. coli* (EAEC) that had acquired an *stx2a* phage, and human illnesses that resulted during 2011 became the largest known HUS outbreak of STEC-related illness in the world [72]. Similarly, an O26:H11 strain 21765, isolated in 2005 during a milk cheese outbreak in France [73] was shown to be an EPEC strain that had probably acquired a *stx2a* phage [27]. In Gonzalez-Escalona et al (2016), the authors demonstrated that some strains of *E. coli* O26:H11 isolated from US cattle were phylogenetically more closely related to ST29 O26:H11 EHECs but t because these did not carry the *stx* phage, they would have been classified as EHEC-like by previous methods [74]. Over the last five years, the analyses of thousands of *E. coli* genomes have revealed that so-called *E. coli* “hybrid strains”–strains that belong to one pathotype but acquire virulence markers, such as *stx* genes, from another pathotype–could be more common than previously believed.

We are heading to a new phase in surveillance of STECs in the US by using a genomic monitoring approach and the genomes of the STEC isolated from FRFDA provides a solid foundation to build upon [75] (https://www.cdc.gov/pulsenet/pathogens/wgs.html). A database already exists that achieves the first goal of source tracking by using core genome information (NCBI pathogen detection tool). However, there is a need for improved databases that allow
for fast analysis of the WGS data for detecting virulence genes, phages and plasmids content, as well as antimicrobial resistance genes.

In conclusion, STECs were isolated from diverse FRFDA food sources during the period study. The contamination frequency was relatively low (median 30 STEC strains isolated per year). However, fifty percent of the STECs analyzed in this study carried either a combination of eae plus stx, or saa plus stx, therefore being potentially pathogenic to humans. Moreover, those STECs carried most of the virulence genes described for STECs causing infections with a diverse range from HC (e.g. ST655 O111:H19 strains) to HUS (e.g. ST21 O26:H11 strains) [20,42]. Some others have not been described as causing disease in humans but have the potential to do so (e.g. ST342 O5:H-unknown strains) since they carried all virulence genes described in pathogenic strains (stx1a, eae-beta1, exhA, tir, and many of the T3SS effectors and non-LEE effectors) (Table 4). Nonetheless, the determination of the presence of STECs in FRFDA with the potential to cause disease in humans reinforces the need to continue surveillance of this important pathogen which is of importance for food safety and public health. Furthermore, the availability of these genomes could provide early warnings of food contamination from cattle or other animals, since some of the STEC isolated were carrying stx2e that have been usually observed causing edema in pigs [76] and are considered as probably non-pathogenic to humans [53]. Here we showed that WGS enabled comparisons across isolates to establish phylogeny, helped in identification of antibiotic resistance by monitoring the presence of antimicrobial resistance genes, and determined the presence of known virulence genes that have been linked with illnesses. A freely accessible dataset of high-quality reference genome sequences of FRFDA was previously unavailable. Future food safety investigations will benefit from the comparisons made possible by this WGS dataset as it allows for the monitoring of the recurrence and emergence of strains in the food supply. It is our goal to help develop a database that will allow for fast source tracking and accurate categorization (low risk or high risk) of STEC food isolates in a more comprehensive manner.

**Supporting information**

S1 Table. Metadata of all *E. coli* isolates analyzed in this study. (XLSX)

S2 Table. *E. coli* virulence genes tested by in silico virulence typing. (XLSX)

S3 Table. *in silico* characterization of STECs for presence of virulence genes and serotype. (XLSX)

S4 Table. ST, serotype and virulence profile of intimin positive non-STEC isolated from FDA regulated foods (2003–2017). (XLSX)

S1 Fig. Minimum spanning tree showing the relationships among all different STECs from this study. The numbers above the connected lines (not to scale) represent allele differences between strains belonging to the same ST. The isolates are colored based on different STs as labeled. (TIF)

**Acknowledgments**

The authors thank Dr. Shanker Reddy for sharing the USDA MDP program food isolates used in this study and the laboratorians within the FDA Office of Regulatory Affairs, Food and Feed
Laboratories who diligently isolated and sequenced many of the isolates here within analyzed. We also want to thank Sabina Lindley for assistance in sequencing some of these *E. coli* strains and Lili Fox Vélez for her helpful comments on this manuscript.

Author Contributions

**Conceptualization:** Narjol González-Escalona, Julie Ann Kase.

**Data curation:** Narjol González-Escalona.

**Formal analysis:** Narjol González-Escalona.

**Investigation:** Narjol González-Escalona, Julie Ann Kase.

**Methodology:** Narjol González-Escalona.

**Software:** Narjol González-Escalona.

**Visualization:** Narjol González-Escalona.

**Writing – original draft:** Narjol González-Escalona, Julie Ann Kase.

**Writing – review & editing:** Narjol González-Escalona, Julie Ann Kase.

References

1. Rangel JM, Sparling PH, Crowe C, Griffin PM, Swerdlow DL. Epidemiology of *Escherichia coli* O157:H7 outbreaks, United States, 1982–2002. Emerg Infect Dis. 2005; 11: 603–609. https://doi.org/10.3201/eid1104.040739 PMID: 15829231

2. Brooks JT, Sowers EG, Wells JG, Greene KD, Griffin PM, Hoekstra RM et al. Non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States, 1983–2002. J Infect Dis. 2005; 192: 1422–1429. JID34662 [pii]; https://doi.org/10.1086/466536 PMID: 16170761

3. 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: 7–15.

4. Beutin L, Krause G, Zimmermann S, Kaulfuss S, Gleier K. Characterization of Shiga toxin-producing *Escherichia coli* strains isolated from human patients in Germany over a 3-year period. J Clin Microbiol. 2004; 42: 1099–1108. https://doi.org/10.1128/JCM.42.3.1099-1108.2004 PMID: 15004060

5. Kuehne A, Bouwknecht M, Havelaar A, Gilsworth A, Hoyer P, Stark K et al. Estimating true incidence of O157 and non-O157 Shiga toxin-producing *Escherichia coli* illness in Germany based on notification data of haemolytic uraemic syndrome. Epidemiol Infect. 2016; 144: 3305–3315. S0950268816001436; https://doi.org/10.1017/S0950268816001436 PMID: 27468812

6. Byrne L, Jenkins C, Launders N, Elson R, Adak GK. The epidemiology, microbiology and clinical impact of Shiga toxin-producing *Escherichia coli* in England, 2009–2012. Epidemiol Infect. 2015; 143: 3475–3487. S0950268815000746; https://doi.org/10.1017/S0950268815000746 PMID: 25920912

7. McAllister LJ, Bent SJ, Petty NK, Skipponing E, Beatson SA, Paton JG et al. Genomic Comparison of Two O111:H- Enterohemorrhagic *Escherichia coli* Isolates from a Historic Hemolytic-Uremic Syndrome Outbreak in Australia. Infect Immun. 2016; 84: 775–781. IAI.01229-15; https://doi.org/10.1128/IAI.01229-15 PMID: 26729762

8. U.S. Department of Agriculture FSIS. Shiga Toxin-Producing *Escherichia coli* in Certain Raw Beef Products. Fed Regist. 2012; 77: 31975–31981.

9. Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. Nat Rev Microbiol. 2004; 2: 123–140. https://doi.org/10.1038/nrmicro818 PMID: 15040260

10. Karmali MA, Mascarenhas M, Shen S, Ziebell K, Johnson S, Reid-Smith R et al. Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. J Clin Microbiol. 2003; 41: 4930–4940. https://doi.org/10.1128/JCM.41.11.4930-4940.2003 PMID: 14605120

11. Nataro JP, Kaper JB. Diarrheagenic *Escherichia coli*. Clin Microbiol Rev. 1998; 11: 142–201. PMID: 9457432

12. Garmendia J, Frankel G, Crepin VF. Enteropathogenic and enterohemorrhagic *Escherichia coli* infections: translocation, translocation, translocation. Infect Immun. 2005; 73: 2573–2585. 73/5/2573; https://doi.org/10.1128/IAI.73.5.2573-2585.2005 PMID: 15845459
13. Khaitan A, Jandhyala DM, Thorpe CM, Ritchie JM, Paton AW. The operon encoding SubAB, a novel cytotoxin, is present in shiga toxin-producing *Escherichia coli* isolates from the United States. J Clin Microbiol. 2007; 45: 1374–1375. JCM.00076-07; https://doi.org/10.1128/JCM.00076-07 PMID: 17287319

14. Paton AW, Srimanote P, Woodrow MC, Paton JC. Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enteroctyce effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. Infect Immun. 2001; 69: 6999–7009. https://doi.org/10.1128/IAI.69.11.6999-7009.2001 PMID: 11598075

15. Garmendia J, Ren Z, Tennant S, Midolli Viera MA, Chong Y, Whale A et al. Distribution of tcpP in clinical enterohemorrhagic and enteropathogenic *Escherichia coli* isolates. J Clin Microbiol. 2005; 43: 5715–5720. 43/11/5715; https://doi.org/10.1128/JCM.43.11.5715-5720.2005 PMID: 16272509

16. Beutin L, Prada J, Zimmermann S, Stephan R, Orskov I, Orskov F. Enterohemolysin, a new type of hemolysin produced by some strains of enteropathogenic E. coli (EPEC). Zentralbl Bakteriol Mikrobiol Hyg A. 1988; 267: 576–588. PMID: 3298285

17. Schmidt H, Kernbach C, Karch H. Analysis of the EHEC hly operon and its location in the physical map of the large plasmid of enterohaemorrhagic *Escherichia coli* O157:h7. Microbiology. 1996; 142 (Pt 4): 907–914.

18. 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: 497–503. PMID: 9966802

19. Schmidt H, Beutin L, Karch H. Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157:H7 strain EDL 933. Infect Immun. 1995; 63: 1055–1061. PMID: 7868227

20. Crowe SJ, Bottichio L, Shade LN, Whitney BM, Corral N, Melius B et al. Shiga Toxin-Producing *E. coli* Infections Associated with Flour. N Engl J Med. 2017; 377: 2036–2043. https://doi.org/10.1056/NEJMoa1615910 PMID: 29166238

21. Erickson MC, Doyle MP. Food as a vehicle for transmission of Shiga toxin-producing *Escherichia coli*. J Food Prot. 2007; 70: 2426–2449. PMID: 17969631

22. Feng PC, Reddy SP. Prevalence and diversity of enterotoxigenic *Escherichia coli* strains in fresh produce. J Food Prot. 2014; 77: 820–823. https://doi.org/10.4315/0362-028X.JFP-13-412 PMID: 24780338

23. Feng PC, Reddy S. Prevalences of Shiga toxin subtypes and selected other virulence factors among Shiga-toxigenic *Escherichia coli* strains isolated from fresh produce. Appl Environ Microbiol. 2013; 79: 6917–6923. AEM.02455-13; https://doi.org/10.1128/AEM.02455-13 PMID: 23995936

24. Allard MW, Bell R, Ferreira CM, Gonzalez-Escalona N, Hoffmann M, Muruvanda T et al. Genomics of the large plasmid of enterohaemorrhagic *Escherichia coli* O157:H7. Microbial Genomics. 2017; 3: 00076-07; https://doi.org/10.1016/j.micgen.2017.11.002 PMID: 29169072

25. Bergholz TM, Moreno Switt AI, Wiedmann M. Omics approaches in food safety: fulfilling the promise? Trends Microbiol. 2014; 22: 275–281. S0966-842 X(14)00018-3; https://doi.org/10.1016/j.tim.2014.01.006 PMID: 24572764

26. Hoffman M, Luo Y, Monday SR, Gonzalez-Escalona N, Ottesen AR, Muruvanda T et al. Tracing Origins of the *Salmonella* Bareilly Strain Causing a Food-borne Outbreak in the United States. J Infect Dis. 2016; 213: 502–508. jvd297 [pii]; https://doi.org/10.1093/infdis/jvd297 PMID: 25995194

27. 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: 3900–3912. AEM.00498-16 [pii]; https://doi.org/10.1128/AEM.00498-16 PMID: 27107118

28. Maiden MC, Harrison OB. The population and functional genomics of the Neisseria revealed with gene-gene approaches. J Clin Microbiol. 2016. JCM.00301-16 [pii]; https://doi.org/10.1128/JCM.00301-16

29. Lorenz SC, Kotewicz ML, Hoffman M, Gonzalez-Escalona N, Fischer M, Kase JA. Complete Genome Sequences of Four Enterohemolysin-Positive (ehxA) Enterocyte Effacement-Negative Shiga Toxin-Producing *Escherichia coli* Strains. Genome Announc. 2016; 4: 4/5/e00846-16 [pii]; https://doi.org/10.1128/genomeA.00846-16 PMID: 27587806

30. Dallman TJ, Byrne L, Lauwers 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: 1672–1680. S0950268814026968 [pii]; https://doi.org/10.1017/S0950268814026968 PMID: 25316375

31. Beutin L, Martin A. Outbreak of Shiga toxin-producing *Escherichia coli* (STEC) O104:H4 infection in Germany causes a paradigm shift with regard to human pathogenicity of STEC strains. J Food Prot. 2012; 75: 408–418. https://doi.org/10.4315/0362-028X.JFP-11-452 PMID: 22289607
32. Pennington H. Escherichia coli O104, Germany 2011. Lancet Infect Dis. 2011. S1473-3099(11)70166-9 [pii]; https://doi.org/10.1016/S1473-3099(11)70166-9 [doi].

33. Li Z, Perez-Osorio A, Wang Y, Eckmann K, Glover WA, Allard MW et al. Whole genome sequencing analyses of Listeria monocytogenes that persisted in a milkshake machine for a year and caused illnesses in Washington State. BMC Microbiol. 2017; 17: 134. https://doi.org/10.1186/s12866-017-1043-1 PMID: 28619007

34. Ogura Y, Ooka T, Iguchi A, Toh H, Asadulghani M, Oshima K et al. Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic Escherichia coli. Proc Natl Acad Sci U S A. 2009; 106: 17939–17944. https://doi.org/10.1073/pnas.0903585106 PMID: 19815525

35. 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: 709–717. https://doi.org/10.1056/NEJMo1106920 PMID: 21737340

36. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH et al. Sex and virulence in Escherichia coli: an evolutionary perspective. Mol Microbiol. 2006; 60: 1136–1151. PMID:172; https://doi.org/10.1111/j.1365-2958.2006.05172.x PMID: 16687971

37. 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: 2410–2426. JCM.00008-15 [pii]; https://doi.org/10.1128/JCM.00008-15 PMID: 25972421

38. 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: 1501–1510. JCM.03617-13 [pii]; https://doi.org/10.1128/JCM.03617-13 PMID: 24574290

39. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. 2012; 67: 2640–2644. dks261 [pii]; https://doi.org/10.1093/jac/dks261 PMID: 22782487

40. Nei M, Tajima F, Tateno Y. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. J Mol Evol. 1983; 19: 153–170. PMID: 6571220

41. Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson’s index of diversity. J Clin Microbiol. 1988; 26: 2456–2466. PMID: 3069867

42. Bielaszewska M, Mellmann A, Bietz 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: 1373–1381. cit055 [pii]; https://doi.org/10.1093/cid/cit055 PMID: 23378282

43. Taylor EV, Nguyen TA, Machesy KD, Koch E, Sotir MJ, Bohm SR et al. Multistate outbreak of Escherichia coli O145 infections associated with romaine lettuce consumption, 2010. J Food Prot. 2013; 76: 939–944. https://doi.org/10.4315/0362-028X.JFP-12-503 PMID: 23726187

44. Watanabe H, Wada A, Inagaki Y, Itoh K, Tamura K. Outbreaks of enterohaemorrhagic Escherichia coli O157:H7 infection by two different genotype strains in Japan, 1996. Lancet. 1996; 348: 939–944. https://doi.org/10.1016/S0140-6736(96)60037-9 PMID: 8614014

45. Bell BP, Goldoft M, Griffin PM, Davis MA, Gordon DC, Tarr PI et al. A multistate outbreak of Escherichia coli O157:H7-associated bloody diarrhea and hemolytic uremic syndrome from hamburgers. The Washington experience. JAMA. 1994; 272: 1349–1353. PMID: 7933395

46. Ethelberg S, Smith B, Torp Dahl M, Lisby M, Boel J, Jensen T et al. Outbreak of non-O157 Shiga toxin-producing Escherichia coli infection from consumption of beef sausage. Clin Infect Dis. 2009; 48: e78–e81. https://doi.org/10.1086/597502 PMID: 19272017

47. Fey PD, Wickert RS, Rupp ME, Safranek TJ, Hinrichs SH. Prevalence of non-O157:H7 shiga toxin-producing Escherichia coli in diarrheal stool samples from Nebraska. Emerg Infect Dis. 2000; 6: 530–533. https://doi.org/10.3201/eid0605.000513 PMID: 10996385

48. Cooper KK, Mandrell RE, Louie JW, Krollach J, Clark TA, Parker CT et al. Comparative genomics of enterohemorrhagic Escherichia coli O145:H28 demonstrates a common evolutionary lineage with Escherichia coli O157:H7. BMC Genomics. 2014; 15: 17. 1471-2164-15-17 [pii]; https://doi.org/10.1186/1471-2164-15-17 PMID: 24410921

49. Feng PC, Delannoy S, Lacher DW, Dos Santos LF, Beutin L, Fach P et al. Genetic diversity and virulence potential of shiga toxin-producing Escherichia coli O111:H21 strains isolated from clinical, environmental, and food sources. Appl Environ Microbiol. 2014; 80: 4757–4763. AEM.01182-14; https:// doi.org/10.1128/AEM.01182-14 PMID: 24889089

50. Mellmann A, Bielaszewska M, Kock R, Friedrich AW, Fruth A, Middendorf B et al. Analysis of collection of hemolytic uremic syndrome-associated enterohemorrhagic Escherichia coli. Emerg Infect Dis. 2008; 14: 1287–1290. https://doi.org/10.3201/eid1408.071082 PMID: 18690698
51. Bielaszewska M, Prager R, Kock R, Mellmann A, Zhang W, Tschape H et al. Shiga toxin gene loss and transfer in vitro and in vivo during enterohemorrhagic Escherichia coli O157 infections in humans. Appl Environ Microbiol. 2007; 73: 3144–3150. AEM.02937-06 [pii]; https://doi.org/10.1128/AEM.02937-06 PMID: 17400784

52. Bielaszewska M, Friedrich AW, Aldick T, Schurk-Bulgirin R, Karch H. Shiga toxin activatable by intestinal mucus in Escherichia coli isolated from humans: predictor for a severe clinical outcome. Clin Infect Dis. 2006; 43: 1160–1167. CID004120; https://doi.org/10.1086/508195 PMID: 17029135

53. Scheutz F, Teel LD, Beutin L, Pierard D, Buvens G, Karch H et al. Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. J Clin Microbiol. 2012; 50: 2951–2963. JCM.00860-12; https://doi.org/10.1128/JCM.00860-12 PMID: 22760050

54. Beutin L, Hammerl JA, Reetz J, Strauch E. Shiga toxin 2A-encoding bacteriophages in enterohaemorrhagic Escherichia coli O104:H4 strains. Emerg Infect Dis. 2014; 20: 1567–1568. https://doi.org/10.3201/ eid2014.131373 PMID: 25152038

55. Melton-Celsa AR, O’Brien AD (1998) Structure, biology, and relative toxicity of Shiga toxin family members for cells and animals. In: Kaper JB, O’Brien AD, editors. Escherichia coli O157:H7 and other Shiga toxin-producing E. coli strains. Washington, DC: ASM Press. pp. 121–128.

56. Shin SW, Shin MK, Jung M, Belayneh KM, Yoo HS. Prevalence of Antimicrobial Resistance and Transfer of Tetracycline Resistance Genes in Escherichia coli isolates from Beef Cattle. Appl Environ Microbiol. 2015; 81: 5560–5566. AEM.01511-15; https://doi.org/10.1128/AEM.01511-15 PMID: 26048929

57. Liu YY, Wang Y, Walsh TR, Yi LX, Zhang R, Spencer J et al. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. Lancet Infect Dis. 2016; 16: 161–168. S1473-3099(15)00424-7; https://doi.org/10.1016/S1473-3099(15)00424-7 PMID: 26603172

58. Karami N, Martner A, Enne VI, Swerkersson S, Adlerberth I, Wold AE. Transfer of an ampicillin resistance gene between two Escherichia coli strains in the bowel microbiota of an infant treated with antibiotics. J Antimicrob Chemother. 2007; 60: 1142–1145. dkm327; https://doi.org/10.1093/jac/dkm327 PMID: 17768176

59. Ruppitsch W, Pietzka A, Prior K, Bletz S, Fernandez HL, Allerberger F et al. Defining and Evaluating a Core Genome Multilocus Sequence Typing Scheme for Whole-Genome Sequence-Based Typing of Listeria monocytogenes. J Clin Microbiol. 2015; 53: 2869–2876. JCM.01193-15 [pii]; https://doi.org/10.1128/JCM.01193-15 PMID: 26135865

60. Chen Y, Gonzalez-Escalona N, Hammack TS, Allard MW, Strain EA, Brown EW. Core Genome Multilocus Sequence Typing for Identification of Globally Distributed Clonal Groups and Differentiation of Outbreak Strains of Listeria monocytogenes. Appl Environ Microbiol. 2016; 82: 6258–6272. AEM.01532-16 [pii]; https://doi.org/10.1128/AEM.01532-16 PMID: 27520821

61. Gonzalez-Escalona N, Gavilan RG, Brown EW, Martinez-Urtaza J. Transoceanic Spreading of Pathogenic Strains of Vibrio parahaemolyticus with Distinctive Genetic Signatures in the recA Gene. PLoS One. 2015; 10: e0117485. https://doi.org/10.1371/journal.pone.0117485 [doi]; PONE-D-14-38928 [pii]. PMID: 25679989

62. Strauss L, Stegger M, Akpaka PE, Alabi A, Breurec S, Coombs G et al. Origin, evolution, and global transmission of community-acquired Staphylococcus aureus ST8. Proc Natl Acad Sci U S A. 2017; 114: E10596–E10604. 1702472114; https://doi.org/10.1073/pnas.1702472114 PMID: 28158405

63. Gonzalez-Escalona N, Jolley KA, Reed E, Martinez-Urtaza J. Defining a Core Genome Multilocus Sequence Typing Scheme for the Global Epidemiology of Vibrio parahaemolyticus. J Clin Microbiol. 2017; 55: 1682–1697. JCM.00227-17; https://doi.org/10.1128/JCM.00227-17 PMID: 28330888

64. Ghanem M, Wang L, Zhang Y, Edwards S, Lu A, Ley D et al. Core Genome Multilocus Sequence Typing: a Standardized Approach for Molecular Typing of Mycoplasma gallisepticum. J Clin Microbiol. 2018; 56. JCM.01145-17; https://doi.org/10.1128/JCM.01145-17 PMID: 29070657

65. Martinez-Urtaza J, van AR, Abanto M, Haendiges J, Myers RA, Trinanes J et al. Genomic Variation and Evolution of Vibrio parahaemolyticus ST36 over the Course of a Transcontinental Epidemic Expansion. MBio. 2017; 8. mBio.01425-17; https://doi.org/10.1128/mBio.01425-17 PMID: 29138301

66. Okello E, Moonens K, Erume J, De GH. Enterotoxigenic Escherichia coli strains are highly prevalent in Ugandan piggeries but disease outbreaks are masked by antibiotic prophylaxis. Trop Anim Health Prod. 2015; 47: 117–122. https://doi.org/10.1007/s11250-014-0694-2 PMID: 25311441

67. Ochoa TJ, Contreras CA. Enteropathogenic Escherichia coli infection in children. Curr Opin Infect Dis. 2011; 24: 478–483. https://doi.org/10.1097/QCO.0b013e32834a8b8b PMID: 21857511

68. Trabulsy LR, Keller R, Gomes TAT. Typical and Atypical Enteropathogenic Escherichia coli. Emerg Infect Dis. 2002; 8: 508–513. https://doi.org/10.3201/eid0805.010385 PMID: 11996687
69. Elliott SJ, Sperandio V, Giron JA, Shin S, Mellies JL, Wainwright L et al. The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic Escherichia coli. Infect Immun. 2000; 68: 6115–6126. PMID: 11035714

70. Karch H, Meyer T, Russmann H, Heesemann J. Frequent loss of Shiga-like toxin genes in clinical isolates of Escherichia coli upon subcultivation. Infect Immun. 1992; 60: 3464–3467. PMID: 1639518

71. McWilliams BD, Torres AG. Enterohemorrhagic Escherichia coli adhesins. Microbiol Spectrum. 2014; 2: EHEC-0003-2013. https://doi.org/10.1128/microbiolspec.EHEC-0003-2013 PMID: 25635238

72. Rohde H, Qin J, Cui Y, Li D, Loman NJ, Hentschke M et al. Open-source genomic analysis of Shiga-toxin-producing E. coli O104:H4. N Engl J Med. 2011; 365: 718–724. https://doi.org/10.1056/NEJMoa1107643 PMID: 21793736

73. Galia W, Mariani-Kurkdjian P, Loukiadis E, Blanquet-Diot S, Leriche F, Brugere H et al. Genome Sequence and Annotation of a Human Infection Isolate of Escherichia coli O26:H11 Involved in a Raw Milk Cheese Outbreak. Genome Announc. 2015; 3. 3/1/e01568-14 [pii]; https://doi.org/10.1128/genomeA.01568-14 PMID: 25700408

74. Bugarel M, Beutin L, Scheutz F, Loukiadis E, Fach P. Identification of genetic markers for differentiation of Shiga toxin-producing, enteropathogenic, and avirulent strains of Escherichia coli O26. Appl Environ Microbiol. 2011; 77: 2275–2281. AEM.02832-10 [pii]; https://doi.org/10.1128/AEM.02832-10 PMID: 21317253

75. Allard MW, Strain E, Melka D, Bunning K, Musser SM, Brown EW et al. Practical Value of Food Pathogen Traceability through Building a Whole-Genome Sequencing Network and Database. J Clin Microbiol. 2016; 54: 1975–1983. JCM.00081-16; https://doi.org/10.1128/JCM.00081-16 PMID: 27008877

76. Oanh TKN, Nguyen VK, De Greve H, Goddeeris BM. Protection of Piglets against Edema Disease by Maternal Immunization with Stx2e Toxoid. Infect Immun. 2012; 80: 469–473. https://doi.org/10.1128/IAI.05539-11 PMID: 22083704