Article

Not peer-reviewed version

Identification and Characterization of ten Escherichia coli Strains encoding Novel Shiga Toxin 2 Subtypes, Stx2n as well as Stx2j, Stx2m, and Stx2o, in the United States

Rebecca L Lindsey *, Arjun Prasad, Michael Feldgarden, Narjol Gonzalez-Escalona, Curtis Kapsak, William Klimke, Angela Melton-Celsa, Peyton Smith, Alexandre Souvorov, Jenny Truong, Flemming Scheutz

Posted Date: 21 September 2023

doi: 10.20944/preprints202309.1481.v1

Keywords: Shiga Toxin-Producing Escherichia coli; Shiga toxin subtype; Stx2n; Genome Sequence

Article

Identification and Characterization of Ten *Escherichia coli* Strains Encoding Novel Shiga Toxin 2 Subtypes, Stx2n as Well as Stx2j, Stx2m, Stx2o, in the United States

Rebecca L Lindsey [1,*], Arjun Prasad [2], Michael Feldgarden [2], Narjol Gonzalez-Escalona [3], Curtis Kapsak [1#,4#,5^], William Klimke [2], Angela Melton-Celsa [6], Peyton Smith [1], Alexandre Souvorov [2], Jenny Truong [7^,8#] and Flemming Scheutz [9]

[1] Enteric Diseases Laboratory Branch, Centers for Disease Control and Prevention, Atlanta, GA, USA, rlindsey1@cdc.gov , koj1@cdc.gov, pjx8@cdc.gov
[2] National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA, aprasad@mail.nih.gov, michael.feldgarden@nih.gov, klimke@ncbi.nlm.nih.gov, souvorov@ncbi.nlm.nih.gov
[3] Center for Food Safety and Applied Nutrition, Food and Drug Administration, College Park, MD, USA, Narjol.Gonzalez-Escalona@fda.hhs.gov
[4] Weems Design Studio, Inc., contractor to Enteric Diseases Laboratory Branch, Centers for Disease Control and Prevention, Atlanta, GA, USA, pjx8@cdc.gov
[5] Theiagen Genomics, Highlands Ranch, CO, USA, kapsakcj@gmail.com
[6] Uniformed Services University of the Health Sciences, Bethesda, MD 20184, angela.melton-celsa@usuhs.edu
[7] Chenega Professional & Technical Services, Chesapeake, VA, USA, odv3@cdc.gov
[8] Oak Ridge Institute for Science and Education, Oak Ridge, TN, USA, odv3@cdc.gov
[9] The International *Escherichia* and *Klebsiella* Centre, Statens Serum Institut, Copenhagen, Denmark, FSC@ssi.dk
* Correspondence: rlindsey1@cdc.gov
^ Current Affiliation:
# Past Affiliation

Abstract: The sharing of genome sequences in online data repositories, allows for large scale analyses of specific genes or gene families. This can result in the detection of novel gene subtypes as well as development of improved detection methods. Here we used publicly available WGS data to detect a novel Stx subtype, Stx2n in two clinical *E. coli* strains isolated in the USA. During this process, additional Stx2 subtypes were detected; six Stx2j one Stx2m strain and one Stx2o, all were analyzed for variability from the originally described subtypes [1,2]. Complete genome sequences were assembled from short or long read sequencing and analyzed for serotype, and ST types. The stx2n and Stx2o WGS were further analyzed for virulence genes pro-phage analysis and phage insertion sites. Nucleotide and amino acid maximum parsimony trees showed expected clustering of the previously described subtypes and a clear separation of the novel Stx2n subtype. WGS data was used to design OMNI PCR primers for the detection of all known Stx1 (283 bp amplicon), Stx2 (400 bp amplicon), *eae* (222bp amplicon) and Stx2f (438 bp amplicon) subtypes. These primers were tested in three different laboratories, using standard reference strains. Analysis of complete genome sequence showed variability in serogroup, virulence genes, ST type, and Stx2 pro-pages showed variability in size, gene composition, and phage insertion sites. The isolates with Stx2n and Stx2o in this study demonstrated toxicity for Vero cells. The Stx2n strain were induced when grown with sub-inhibitory concentrations of Cip, and toxicity was detected. Taken together, these data highlight the need to reinforce genomic surveillance to identify the emergence of potential new Stx2 or Stx1 variants. The importance of this surveillance has a paramount impact in public health. Per our description in this study, we suggest that 2017C-4317 be designated as the Stx2n type-strain.

 doi:10.20944/preprints202309.1481.v1

Keywords: Shiga Toxin-Producing *Escherichia coli*; Shiga toxin subtype; Stx2n; genome sequence

1. Introduction

Shiga toxin (Stx) is the defining virulence factor in Shiga Toxin-Producing *Escherichia coli* (STEC) which can cause gastrointestinal illness with possible life-threatening complications in humans. Two major types of Stx, Stx1 and Stx2, are further divided into subtypes, Stx1 (a, c, d) and 14 Stx2 (a-m, o).

A 2020 study from the EU found that O157 STEC was only isolated in 20.6 % of the confirmed cases of human STEC infections, the remaining 79.4% of confirmed cases were associated with non-O157 STEC [3]. Therefore, methods for detection of toxin genes instead of or in addition to serogroup detection are important diagnostic tools for STEC infection.

In 2012, a standardized Stx nomenclature was established for Stx1/Stx2 and associated subtypes, which included Stx2a-Stx2g [4]. In the last ten years, additional Stx2 subtypes, Stx2h-Stx2m have been described and named following the standardized Stx nomenclature: Stx2h [5], Stx2i [6], Stx2j [1], Stx2k [7], Stx2l [8], Stx2m [2], and Stx2o [1].

As part of evaluation of the AMRFinderPlus tool [9] for the detection of *stx* variants in the Pathogen Detection system, we found 116 (0.2%) genomes among over 60,000 *E. coli* and *Shigella* genomes screened with stx_2 B subunit sequences that fell just below the cutoffs for stx_2 subtypes a through g. Several of the new variants identified were in the process of being characterized and published by other groups (e.g., Stx2j, Stx2m and Stx2o). Among these 116 genomes, two were positive for a novel stx_2 subtype, provisionally designated stx_{2n}. Ten isolates; six Stx2j, one Stx2m, two Stx2n, and one Stx2o, were selected for further analysis and characterization for this study.

2. Materials and Methods

2.1. Detection of Novel Stx2 subtypes

Over 60,000 *Escherichia coli* and *Shigella* isolates with short-read data included in the National Center for Biotechnology Information (NCBI) Pathogen Detection System (https://www.ncbi.nlm.nih.gov/pathogens) as of February 2, 2020 were screened by both *de novo* assembly using SKESA [10] and targeted assembly using SAUTE with characterized *stx* sequences as targets [11]. Resulting assemblies were analyzed with AMRFinderPlus which includes curated Hidden Markov models (HMMs) that can identify novel divergent *stx* genes [9].

2.2. Collection of STEC strains

The standard operating procedure (SOP) for the PulseNet USA, the molecular surveillance network for foodborne disease in the United States [12] includes upload of raw sequence reads to NCBI, where the novel *stx* subtypes were detected. Limited metadata is available within the NCBI BioSample records for these clinical isolates due to laws that prevent sharing of personally identifiable information (PII) [13].

2.3. Illumina and Oxford Nanopore Sequencing and Assembly

DNA was extracted from bacterial cells using the Promega Wizard kit (Promega Wizard Genomic DNA Purification Kit, Promega Corporation, Madison, WI), wide-bore pipette tips and minimal handling were used to produce high molecular weight DNA. A single DNA extract was used for all sequencing methods. Illumina MiSeq libraries were prepared with the DNA Prep Library kit (Illumina, Inc., San Diego, CA), using modified bead ratios for optimal fragment size, following the PulseNet standard operating procedure (SOP) PNL35 (https://www.cdc.gov/pulsenet/pathogens/wgs.html) and sequenced to a minimum of 40X coverage

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2023 doi:10.20944/preprints202309.1481.v1

[14]. Nanopore MinION libraries were prepared with the Rapid Barcoding kit according to manufacturer's protocol without size selection or normalization and sequenced for 72 hours on the R9.4.1 flow cell (Oxford Nanopore Technologies, Oxford, UK).

Illumina raw reads were analyzed with a PulseNet customized version of BioNumerics (Applied Maths, Sint-Martens-Latem, Belgium), a commercial off-the-shelf data analysis and management software to assemble (SPAdes v2.2) then analyze the WGS [12].

The complete hybrid genome assembly comprising the chromosome and plasmid(s) for the Stx2n and Stx2o isolates was obtained by *de novo* assembly using nanopore data and Flye v2.8 [15]. Second Hybrid genome assembly was created using both Illumina and Nanopore reads with Unicycler v0.4.8 [16]. The hybrid and Flye assemblies for each isolate were aligned with Mauve v2.4.0 [17] to look for any disagreement in synteny, size, or completeness. Since both genome assemblies (hybrid and nanopore only) for these three isolates agreed in all those requirements, the hybrid assembly was determined to be the final assembly (*i.e.*, complete genome). Unicycler assembled the chromosome and plasmids as circular closed and oriented the chromosome to start at the *dnaA* gene. The genomes were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP v5.0, http://www.ncbi.nlm.nih.gov/genome/annotation_prok) [18].

2.4. WGS based Characterization

The serotype and virulence gene content of the Stx2j, Stx2m, Stx2n and Stx2o assemblies were identified using the Center for Genomic Epidemiology (http://www.genomicepidemiology.org) web-based Serotype Finder 2.0.2 and Virulence Finder 1.5 tools [19–21].

Multi-Locus Sequence Types were confirmed using an *in silico E. coli* MLST approach, based on the information available at the *E. coli* MLST website (Enterobase https://enterobase.warwick.ac.uk/species/index/ecoli) and using Ridom SeqSphere+ software v2.4.0 (Ridom; Münster, Germany) (http://www.ridom.com/seqsphere), through BioNumerics (using the same scheme), or with Torsten Seemann's command line software mlst v2.23.0 (https://github.com/tseemann/mlst#cit https://github.com/tseemann/mlst#citations ations). Seven housekeeping genes (adk, fumC, gyrB, icd, mdh, purA, and recA), described previously for *E. coli* [22], were used for MLST analysis and to assign numbers for alleles and sequence type (ST) (https://github.com/tseemann/mlst/tree/master/db/pubmlst/ecoli_achtman_4).

2.5. Stx subtyping

The Stx subtypes of the selected STEC isolates were determined by ABRicate version 0.8.10 (https://github.com/tseemann/abricate, accessed on 1 March 2020) with the default parameters. Briefly, a stx subtyping database was created with ABRicate by including representative nucleotide sequences of all identified Stx1 and Stx2 subtypes.

The assemblies were then searched against the *stx* subtyping database. For the *stx* genes that yield an identity below 96% with the nearest known stx subtype, the full nucleotide sequences were extracted and compared to the GenBank database with the NCBI Blast tool.

The representative nucleotide sequences of all the *stx*$_2$ subtypes and variants (*stx2a-stx2m*, *stx2o*) described previously were downloaded from GenBank (Stx2h [5], Stx2i [6], Stx2j [1], Stx2k [7], Stx2l [8], Stx2m [2], Stx2o [1]). The amino acid sequences for the combined A and B subunits of Stx2 holotoxin were translated from the open reading frames. The full amino acid and nucleotide sequences were aligned to calculate the genetic distances between *stx*$_2$/Stx2 sequences. Evolutionary unrooted trees were created from maximum parsimony cluster analysis using 100 bootstrap resamples. Also, the amino acid sequences were analyzed for sequence motifs that support the phylogenetic analyses using BioNumerics version 7.6 (Applied Maths, Ghent, Belgium), as previously described Scheutz, F., et al., [4].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2023 doi:10.20944/preprints202309.1481.v1

2.6. Cytotoxicity, Ciprofloxacin (Cip) induction, and activation assays

The level of cytotoxicity from culture supernatant fractions or cell fractions were determined on Vero cells as previously described [23]. Sub-lethal concentrations of ciprofloxacin (5ng/mL) were added to some cultures for evaluation of induction. The toxins were tested for activation by incubation of the supernatant fractions with mouse intestinal mucus or a buffer control for 1-2 hours then determining the toxicity on Vero cells as described previously [24].

2.7. Polymerase Chain Reaction (PCR) primers to detect all described stx1 and stx2 subtypes

Primers for this study, were redesigned and tested independently in three laboratories for detection of all 18 Stx1/Stx2 subtypes (Table S1). PCR was done for each target in a total volume of 20 µl with 10 µl HotStarTaq Master Mix Kit (Qiagen), 5 µl of primer mix (20 µM each primer), and 5 µl supernatant of boiled lysate. The thermocycler conditions were 95°C for 15 min followed by 35 cycles of 94°C for 50 s, 62°C for 40 s, and 72°C for 50 s, ending with 72°C for 3 min. PCR amplicons were stored at 4°C. Amplicons were separated on a 2% agarose gel stained with GelRed for a total of 30 minutes at 100 volts.

Table 1. Primers for detection of all known Stx1 and Stx2 subtypes.

Primer Name	Primer Sequence	Amplicon Size (bp)	Reference
stx2-PS8-F	5′-TCAC**Y**GGTTTCATCATATCTGG	400	This study
stx2-PS7-R	5′-GCCTGTC**B**CCA**S**TTATCTGACA		
PS19 *stx2f*-F	5′-GTACAGGGATGCAGATTGGGCG	438	This study
PS20 *stx2f* -R	5′-CTTTAATGGCCGCCCTGTCTCC		
PS17 *eae*-F	5′-CGGCTATTTCCGCATGAGCGG	223	This study
PS18 *eae*-R-NEW	5′AGTT**D**ACACCAAY**W**GTC**R**CCGC		
stx1 F3b	5′-CTGATGATTGATAGTGGCACAGG	283	This study
stx1 OMNI-R1	5′-GCGATTTATCTGCATCCCCGTAC		

2.8. Stx2n and Stx2o pro-phages annotation and discovery

The pro-phages carrying stx_{2n} and stx_{2o} genes were identified using Phaster (https://phaster.ca). The Stx2 pro-phage region flanked by *attL* and *attR* sites from each genome strain were extracted and annotated using Galaxy tracker Prokka 1.14.6 [25] and visualized with SnapGene Viewer v6 (https://www.snapgene.com/snapgene-viewer).

2.9. Data availability

Raw sequences, along with their limited metadata, are publicly available in the sequence read archive (SRA) housed by the National Center for Biotechnology Information (NCBI) under BioProject PRJNA218110, the accession numbers are shown in Table 3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2023 doi:10.20944/preprints202309.1481.v1

3. Results

3.1. Identification of two Novel Stx2 Subtypes

An evaluation of the AMRFinderPlus tool for the detection of stx gene variants in the Pathogen Detection system, found 116 (0.2%) genomes among over 60,000 E. coli and Shigella genomes screened with stx2 B subunit sequences that fell just below the cutoffs for stx2 subtypes a through g. Several of the new variants identified were in the process of being characterized and published by other groups (e.g., Stx2j, Stx2m and Stx2o). Among these 116 genomes, two were positive for a novel Stx2 subtype, provisionally designated Stx2n, and the remaining 114 were closely related to recently described identical subtypes Stx2j [MZ229608 and MZ571121] or Stx2o [MZ229604], or Stx2m [OQ054797]. Three strains 2013C-3244, 2017C-4317 and 2018C-3367 were selected for resequencing to generate closed genomes for further analysis.

The in-house stx-subtyping based on whole-genome sequences showed that the stx2 sequences from a representative Stx2n strain shared less than 94.6% nucleic acid sequence identities with other stx2 subtypes. stx2 genes and Stx2 proteins were extracted from the genome assemblies and compared against the GenBank database using NCBI BLAST. These comparisons showed the highest similarity (94.6%) with the Stx2n strain. When comparing sequences of Stx2 holotoxin, Stx2n shared 72.2 to 94.6% similarity with the other 14 described Stx2 subtypes at nucleic acid level and 83.9 to 95% at amino acid level (Table 2). The last six amino acids in the A subunit were absent, which is also where the amino acid differences between the seven variants of Stx2j were found. All seven Stx2j variants had identical B subunits. These results suggest that the two provincial STEC strains harbor novel Stx2 subtypes. Based on the standardized nomenclature for Stx2 [4], the new Stx2 subtype was designated Stx2n.

Table 2. Nucleotide\amino acid identities (%) using Neighbor Joining comparison in BioNumerics version 8.1 (Applied Maths, Biomérieux), between Stx2n, Stx2o and representatives of other described Stx2 subtypes. 1. Stx2a (EDL933, X07865), 2. Stx2b (EH250, AF043627), 3. Stx2c (031, L11079), 4. Stx2d (C165-02, DQ059012), 5. Stx2e (S1191 (M21534), 6. Stx2f (F08-101-31, AB472687), 7. Stx2g (7v, AY286000), 8. Stx2h (STEC299, CP022279), 9. Stx2i (CB10366, FN252457), 10. Stx2j (5447, MZ571121), 11. Stx2k (STEC309, CP041435), 12. Stx2l (FHI 1106-1092, AM904726), 13. stx2m (2001F31428, OQ054797) 14. stx2n 2017C-4317 (GCA_013342905.2) and 15. stx2o (03-3638, MZ229604). Bold values highlight the sequence identities of identified Stx2n and Stx2o subtypes with previously reported Stx2 subtypes.

Nucleotide\ Amino acid	1	2	3	4	5	6	7	8	9	10	11	12	13	**14**	**15**
	stx2 a	*stx2 b*	*stx2 c*	*stx2 d*	*stx2 e*	*stx2 f*	*stx2 g*	*stx2 h*	*stx2 i*	*stx2 j*	*stx2 k*	*stx2 l*	*stx2 m*	***stx2 n***	***stx2 o***
1. Stx2a		91,9	98,4	96,9	92,2	70,8	94,2	91,7	93,2	89,5	94,4	95,5	93,2	**87,9**	**89,7**
2. Stx2b	95,3		92,2	93,4	89,4	70,6	91,4	92,2	89,2	89,2	91,3	90,0	90,6	**87,8**	**90,2**
3. Stx2c	99,2	95,4		97,4	91,9	70,5	93,1	91,7	92,4	89,6	94,7	94,8	92,2	**88,1**	**89,7**
4. Stx2d	98,4	96,0	98,5		92,1	70,5	94,0	92,2	92,9	90,2	96,1	94,8	91,5	**88,2**	**90,2**
5. Stx2e	95,2	94,2	94,9	95,7		74,8	92,2	90,2	94,7	88,3	93,8	95,0	88,9	**86,9**	**88,8**
6. Stx2f	81,9	81,2	81,4	82,0	84,1		71,4	71,2	71,3	71,1	71,0	71,5	71,1	**72,2**	**70,4**
7. Stx2g	97,0	95,1	96,5	97,1	95,9	82,2		91,9	94,5	88,2	92,9	92,8	91,2	**87,0**	**89,1**
8. Stx2h	95,2	95,3	95,0	95,6	94,4	81,8	95,5		92,0	92,1	92,8	91,1	92,1	**91,4**	**94,1**
9. Stx2i	95,9	93,7	95,4	96,1	97,3	82,3	96,5	95,4		88,6	96,5	95,1	90,0	**88,1**	**89,4**
10. Stx2j	93,5	93,2	93,7	94,2	92,6	82,7	92,7	93,5	92,5		90,6	89,8	89,2	**88,1**	**90,9**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2023 doi:10.20944/preprints202309.1481.v1

11.Stx2k	97,2	95,4	97,3	98,2	97,2	82,2	96,8	95,9	97,9	94,1		95,5	90,9	**88,1**	**90,3**
12.Stx2l	96,8	94,1	96,4	97,1	97,3	82,3	96,2	94,7	97,3	93,9	97,8		89,9	**87,4**	**89,9**
13.Stx2m	95,7	95,1	95,3	95,5	94,3	82,0	96,1	94,8	94,5	92,1	95,0	93,8		**87,7**	**88,9**
14.Stx2n	**93,2**	**93,1**	**93,4**	**93,6**	**92,6**	**83,9**	**93,5**	**95,0**	**93,1**	**91,8**	**93,9**	**92,4**	**93,2**		**91,4**
15.Stx2o	**94,3**	**94,0**	**94,5**	**94,7**	**93,6**	**81,8**	**94,5**	**96,9**	**94,3**	**92,8**	**95,2**	**93,8**	**93,8**	**94,6**	

The Stx2 subtype amino acid comparison using a maximum parsimony tree (Figure 1) included the novel strains described here: 2017C-4317, 2013C-3244, 2018C-3367, 03-3638 [5]. Strains 2012C-4221*, 2019C-4307, 2019C-4332, 2010C-4332, PNUSAE011983 and 2019C-3762 from this study grouped with previously described Stx2j and Stx2m sequences. Supplementary Table S2 list data for all the sequences in Figures 1 and 2. *O162 was determined by in silico serotyping as positive for both wzmO162 and wztO162. Phenotypic O grouping was positive for O101.

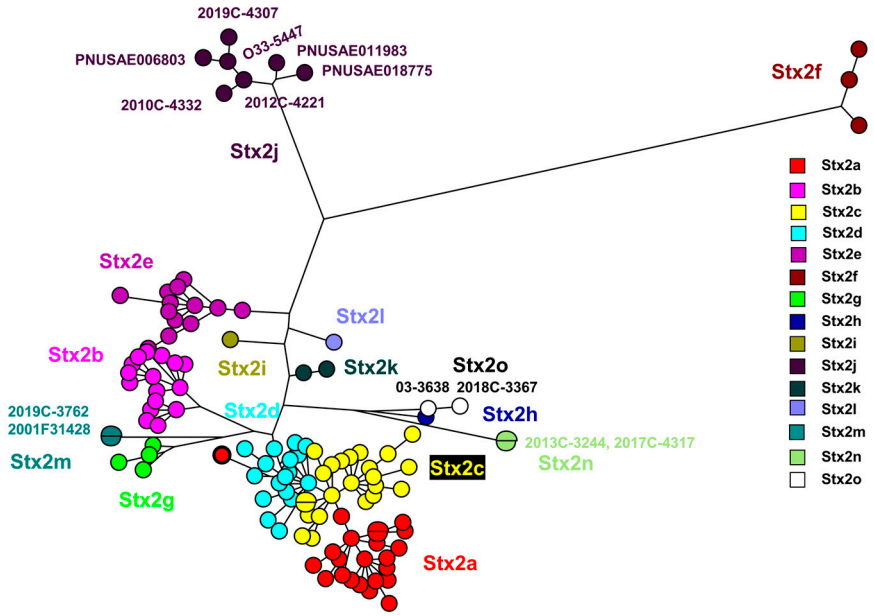

Figure 1. Stx2 subtype amino acid comparison of 117 sequences using maximum parsimony tree. Stx2a, Stx2b, Stx2c, Stx2d, Stx2e, Stx2f, Stx2g adapted from [4]. Stx2h [5], Stx2i [6], Stx2j [1] and this study, Stx2k [7,26], Stx2l [8] and Stx2m [2] and this study, Stx2n (this study) and Stx2o [1] and this study). The Stx2a outlier (Stx2a-08-BMH-17-0026, Acc. No. MZ229605, circled in black) has an "EDD" motif in the B subunit, and is therefore defined as Stx2a, see Gill et al. (5). Supplementary table S2 lists all the sequence information.

The Stx2 subtype nucleotide comparison using a maximum parsimony tree (Figure 2) included the novel strains described in this study: 2017C-4317, 2013C-3244, 2018C-3367, and previously mentioned 03-3638 [5]. Strains 2012C-4221, 2019C-4307, 2010C-4332, PNUSAE018775, PNUSAE006803, PNUSAE011983 and 2019C-3762 , from this study grouped with previously described Stx2j and Stx2m sequences [2]. These strains were included in primer design to ensure detection of known diversity present in Stx2j and Stx2m.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2023 doi:10.20944/preprints202309.1481.v1

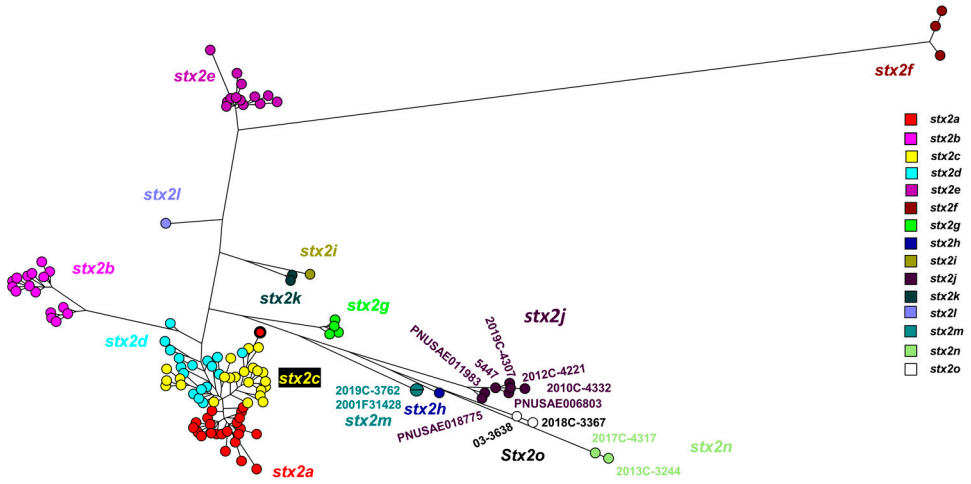

Figure 2. stx2 subtype nucleotide comparison of 117 sequences using maximum parsimony tree. stx2a, stx2b, stx2c, stx2d, stx2e, stx2f, stx2g adapted from [4]. stx2h [5], stx2i [6], stx2j [1], stx2k [7,26], stx2l [8], stx2m [2] and this study and stx2n (this study) and stx2o [1] and this study. The stx2a outlier (Stx2a-08-BMH-17-0026, Acc. No. MZ229605) in Figure 1 is circled in black. Supplementary table S2 lists all the sequence information.

3.2. WGS based Characterization of Stx2n and a Stx2o strains

The ten strains in this study are all different serotypes and STs (Table 3) and show variability across time (isolated from 2010-2019).

Table 3. Characterization of Stx2j, Stx2m, Stx2n and a Stx2o-producing STEC strains.

Stx Subtype	CDC isolate ID	BioSample	ST type	O:(K):H type	PNID	Assembly
stx2i	2010C-	SAMN04377	5662	O158:H23	PNUSAE001	GCA 012764415
stx2i	2012C-	SAMN08579	5350	O162:H6	None	GCA 003018235
stx2i	2019C-	SAMN12361	5736	O32:K87:	PNUSAE027	GCA 011901845
stx2i	See	SAMN10170	491	ONT:H45	PNUSAE018	GCA 003903075
stx2i	See	SAMN07177	5923	O33:H14	PNUSAE006	GCA 012463025
stx2i	See	SAMN08595	657	O183:	PNUSAE011	GCA 012253565
stx2m	2019C-	SAMN11569	9312	O38:H39	PNUSAE024	GCA 011950125
Stx2n*	2013C-	SAMN04578	1385	O1:K22:H	None	GCA 012711215
Stx2n*	2017C-	SAMN07709	70	O23:H15	PNUSAE009	GCA 013342905
Stx2o	2018C-	SAMN08799	80	O75:H7	PNUSAE012	GCA 012224845

*Novel stx subtypes described in this study. [1] Both lactose positive and negative colonies were found. [2] Closed hybrid assemblies generated for this study.

WGS virulence analysis showed that thirty-two virulence genes were harbored between the three isolates, the number of virulence genes varied between the three isolates with 2017C-4317 harboring 9 of 32, 2013C-3244 harboring 18 of 32, and 2018C-3367 harboring 27 of 32 total virulence genes (Table 4). All three isolates harbored stx2, chuA, gad, kpsE, kpsMII_K5, sitA and terC (Table 4). In addition, 2017C-4317 harbored virulence genes; stx2n, traT, and eilA. 2013C-3244, harbored additional virulence genes; stx2n, focC, fyuA, iroN, irp2, iss, ompT, sfaD, sfaS, tcpC, vat, and yfcV. The virulence profile for 2018C-3367 harbored virulence genes; stx2o, focC, fyuA, iroN, irp2, iss, ompT, sfaD, vat, yfcV, neuC, clbB, cnf1, hra, ibeA, mchB, mchC, mchF, mcmA, and pic (Table 4). The presence of fyuA, vat, and yfcV qualifies 2013C-3244 and 2018C-3367 as UPEC$_{HM}$ according to the current definition of UPEC [27].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2023 doi:10.20944/preprints202309.1481.v1

Table 4. Presence of virulence genes in the two Stx2n and a Stx2o-producing STEC sequences.

Virulence gene	Function	2013C-3244	2017C-4317	2018C-3367
stx2n	Shiga toxin 2	+	+	-
stx2o	Shiga toxin 2	-	-	+
chuA	Outer membrane hemin receptor	+	+	+
focC	S fimbrial/F1C minor subunit	+	-	+
*fyuA**	Siderophore receptor	+	-	+
gad	Glutamate decarboxylase	+	+	+
iroN	Enterobactin siderophore receptor protein	+	-	+
irp2	High molecular weight protein 2 non-ribosomal peptide synthetase	+	-	+
iss	Increased serum survival	+	-	+
kpsE	Capsule polysaccharide export inner-membrane protein	+	+	+
kpsMII_K5	Polysialic acid transport protein; Group 2 capsule	+	+	+
ompT	Outer membrane protease (protein protease 7)	+	-	+
sfaD	S fimbrial/F1C minor subunit	+	-	+
sfaS	S-fimbriae minor subunit	+	-	-
sitA	Iron transport protein	+	+	+
tcpC	Tir domain-containing protein	+	-	-
terC	Tellurium ion resistance protein	+	+	+
*vat**	Vacuolating autotransporter toxin	+	-	+
*yfcV**	Fimbrial protein	+	-	+
traT	Outer membrane protein complement resistance	-	+	-
eilA	Salmonella HilA homolog	-	+	-
neuC	Polysialic acid capsule biosynthesis protein	-	-	+
clbB	Hybrid non-ribosomal peptide / polyketide megasynthase	-	-	+
cnf1	Cytotoxic necrotizing factor	-	-	+
hra	Heat-resistant agglutinin	-	-	+
ibeA	Invasin of brain endothelial cells	-	-	+
mchB	Microcin H47 part of colicin H	-	-	+
mchC	MchC protein	-	-	+
mchF	ABC transporter protein MchF	-	-	+
mcmA	Microcin M part of colicin H	-	-	+

 doi:10.20944/preprints202309.1481.v1

pic	serine protease autotransporters of Enterobacteriaceae (SPATE)	-	-	+
usp	Uropathogenic specific protein	-	-	+
Pathotype		STEC/UPEC$_{HM}$	STEC	STEC/UPEC$_{HM}$

* Genes that qualify strains as UPEC$_{HM}$ according to current definition of UPEC (27).

3.3. Detection of Shiga Toxin Production

Culture supernatant from each strain was tested for cytotoxicity on Vero cells as described previously [23]. The isolates in this study demonstrated toxicity for Vero cells, Table 5. Additionally, strain 2012C-4221 could be induced when grown with sub-inhibitory concentrations of ciprofloxacin, and toxicity was detected, Table 5. All the strains were positive for stx2 when tested with OMNI primers, described in this publication (Table 5). None of the supernatants showed increased toxicity after incubation with mouse intestinal mucus (activation), although the control Stx2d sample did show activation (data not shown).

Table 5. Results of Vero Cell Assay.

Stx Subtype	CDC isolate ID	Log CD_{50}/mL supernatant	PCR confirmation	ciprofloxacin induction
stx2j	2010C-4332	4.6	+	No
stx2j	2012C-4221	4.5	+	Yes#
stx2j	2019C-4307	3.4	+	No
stx2m	2019C-3762	3.2	+	No
Stx2n	2013C-3244	3.2	+	No
Stx2n	2017C-4317	2.0^	+	No
Stx2o	2018C-3367	3.0	+	No

#Toxin is detectable after the strain is grown in the presence of ciprofloxacin. ^cell-associated toxicity.

3.4. Design and testing of new OMNI PCR primers

The original stx1 primer to detect all stx1 subtypes [4] was redesigned to produce a slightly larger fragment by using a conserved upstream sequence as the new forward primer, now identified as stx1 F3b and by reversing the original forward primer (stx1-det-F1), now stx1 OMNI-R1 (table 1). This change allows for the detection of a slightly larger fragment (283 bp) and a clear separation of the stx1 fragment from the newly developed fragment for the detection eae. The stx2 primers (stx2-PS8-F, stx2-PS7-R) were redesigned to detect all known stx2 subtypes stx2a-stx2o with a 399bp amplicon, except for stx2f. It needs the stx2f specific primers for a 438bp amplicon (This study, table 1). The eae primers (PS17 eae-F, PS18 eae-R-NEW) are designed to detect all known eae genes, this includes all variants of eae found in Citrobacter spp. These primers were tested in three different laboratories, using standard reference strains to confirm detection and amplicon size.

3.5. Identification of the Stx2n and Stx2o pro-phages in the strains from this study

The Stx2 pro-phages were identified in different locations for each strain from this study (Supplementary figure 1). The Stx2 pro-phages were all different sizes and gene compositions, even among the two stx2n strains (Figure 3). stx2n pro-phage in strain 2013C-3244 was 45.2Kb in size and the %GC content was 49.97. stx2n pro-phage in strain 2017C-4317 was 75.1Kb in size and the %GC content was 50.76. stx2o pro-phage in strain 2018C-3367 was 45.9Kb in size and the %GC content was 49.63. The attL and attR were also different for each Stx2 pro-phages. For strain 2013C-3344 the attL sequence for the stx2n pro-phage was TGGCGAAAAACTG and located at the following coordinates

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 September 2023 doi:10.20944/preprints202309.1481.v1

in the chromosome 2,048,352, while the identical attR sequence was located at 2,086,928. In the case of the stx2n for strain 2017C-4317 the attL sequence for the stx2n pro-phage was TTAATTAATTTA and located at the following coordinates in the chromosome 2,907,157, while the identical attR sequence was located at 2,982,293. For strain 2018C-3367 the attL sequence was TCAATCACTTACA and located at 2,071,326, while the identical attR was located at 2,113,169. Most of the genes identified in these three pro-phages coded for hypothetical proteins.

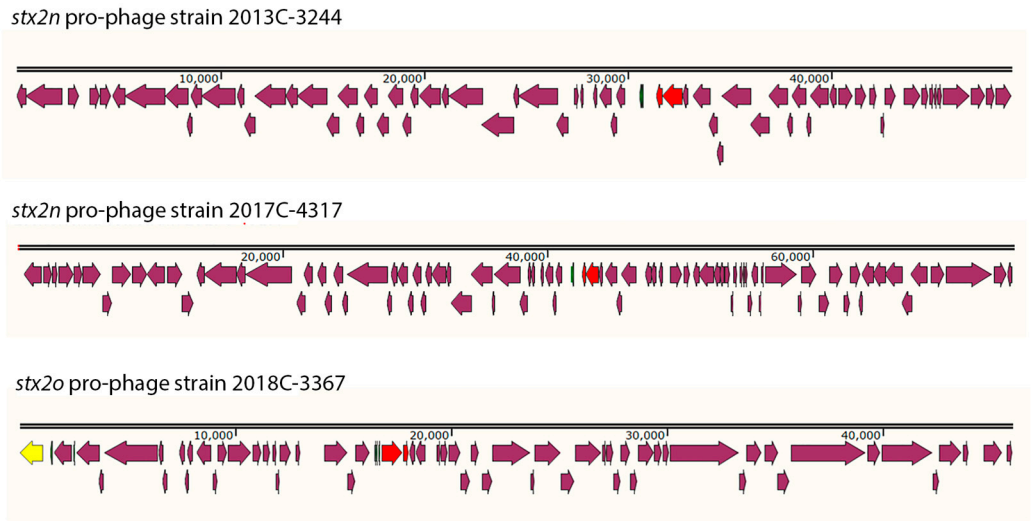

Figure 3. Schematic representation of the stx2n or stx2o pro-phages found in the three different strains. The pro-phages are different sizes and not draw to scale. The stx2 gene A and B units are in red and the integrase is colored yellow.

4. Discussion

In our present study, novel Stx-producing STEC strains were isolated from patients in clinical settings in the United States. Six *stx2j* strains were included in figures 1, 2 and table 3, to demonstrate the diversity in the *stx2j* subtype when compared to each other and strains described by Gill *et.al.* Here, *Stx2j* subtype strains, 2010C-4332, 2012C-4221, 2019C-4307, PNUSAE018775, PNUSAE006803 and PNUSAE011983 were identified in six different serogroups (O158, O162, O32, ONT, O33 and O183) and STs (5662, 5350, 5736, 491, 5923 and 657) over a period of 11 years. The *stx2m* strain, 2019C-3762, was included here to demonstrate the diversity in this subtype when compared to strains described by Bai *et.al.* The two Stx2n-STEC isolates, 2017C-4317 and 2013C-3244, show diversity in serotype (O23 and O1) and ST (70 and 1385). The Stx2o strain, 2018C-3367, was included here to demonstrate the diversity in this subtype when compared to strains described by *Gill, et al* [1]. The genome assemblies for the Stx2 subtype strains (Table 3) will be published in future Microbial Resource Announcements.

The stx2 pro-phages for Stx2n and Stx2o strains described here (2013C-3244, 2017C-4317, 2018C-3367) were identified in different locations (figure 3), were all different sizes and gene compositions (Table 4), even among the two *stx2n* strains (figure S1). The facts that Our findings that the strains carrying these stx_2 subtypes have different predicted serogroups, are in separate ST classes, and were isolated in different years and different locations demonstrates that the phages that encode these toxin subtypes are mobile and have spread among different *E. coli*, as has been shown for other *stx*-phages [28].

The virulence gene profile (Table 3) highlights the variability of known virulence genes among the three isolates, with 2017C-4317 harboring 9 of 32, 2013C-3244 harboring 18 of 32, and 2018C-3367 harboring 27 of 32 total virulence genes (Table 4). Of note the STEC strains 2013C-3244 and 2018C-3367, carry the genes *fyuA, vat,* and *yfcV*, which qualifies these two strains as UPEC$_{HM}$ according to current definition of UPEC [27]. Strains classified as multiple pathotypes can be more dangerous to human health because once an initial pathotype is detected, analysis may stop, missing additional

virulence genes related to a second pathotype related to human illness. The O104 outbreak in Europe was caused by a strain that was both STEC and EAEC [29]. A complete WGS analysis of *E. coli* includes databases from CGE that were designed to provide a complete examination of important genes such as those for serotyping, virulence genes and pathotype [19,27]. We note that the public contribution of surveillance data by groups such as PulseNet and publicly available analysis results such as MicroBIGG-E (https://www.ncbi.nlm.nih.gov/microbigge) demonstrate the power of large-scale and open data analysis to identify novel genes and variants important to public health.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Table S1: Supplementary Table S1; Statens Serum Institut strain collection of reference strains harboring the *stx* gene subtypes, their O:H serotype, additional virulence genes and identification numbers. Table S2: Supplementary Table S2: Sequence information including Accession numbers, protein and nucleotide sequences. Figure S1 Location of pro-phages identified by Phaster (https://phaster.ca) in the chromosome of the strains carrying novel stx2 types (stx2n and stx2o) reported in this study.

Author Contributions: Conceptualization, A.P., B.K., F.S.C., and R.L.; methodology, M.F., A.S., and A.P.; validation, A.M-C., C.K., N.G-E., and M.F.; investigation, A.M-C., F.S.C., J.T., C.K., N.G-E., R.L., and P.S.; writing—original draft preparation, A.P., N.G-E, and R.L.; writing—review and editing, A.P., B. K., and F.S.C.; visualization, F.S.C. and N.G-E; project administration, A.P. and R.L.

Funding: For CDC associated work, this research was funded by the Advanced Molecular Detection (AMD) Initiative grant number AMD-21 at the Centers for Disease Control and Prevention. This project was supported in part by an appointment to the Research Participation Program at the Centers for Disease Control and Prevention administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the Centers for Disease Control and Prevention. As well as NCBI associated research was supported by the National Center for Biotechnology Information of the National Library of Medicine (NLM) and the National Institute of Allergy and Infectious disease (NIAID), National Institutes of Health. Dr. Flemming Scheutz was partially funded by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 773830.

Data Availability Statement: The whole genome sequence data for this study is publicly available at NCBI. Specific NCBI Accession numbers are listed in Table 3 and in supplementary tables S1 and S2. The findings and conclusions in this report are those of the author(s) and do not and do not reflect the view of the Centers for Disease Control and Prevention, the Department of Health and Human Services, or the United States government. Furthermore, the use of any product names, trade names, images, or commercial sources is for identification purposes only, and does not imply endorsement or government sanction by the U.S. Department of Health and Human Services.

Conflicts of Interest: The authors declare no conflict of interest. The funders and contract company, Weems Design Studio, had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the result.

References

1. Gill, A.; Dussault, F.; McMahon, T.; Petronella, N.; Wang, X.; Cebelinski, E.; Scheutz, F.; Weedmark, K.; Blais, B.; Carrillo, C. Characterization of Atypical Shiga Toxin Gene Sequences and Description of Stx2j, a New Subtype. J Clin Microbiol **2022**, 60, e0222921, doi:10.1128/jcm.02229-21.
2. Bai, X.; Scheutz, F.; Dahlgren, H.M.; Hedenstrom, I.; Jernberg, C. Characterization of Clinical Escherichia coli Strains Producing a Novel Shiga Toxin 2 Subtype in Sweden and Denmark. Microorganisms **2021**, 9, doi:10.3390/microorganisms9112374.
3. European Food Safety, A.; European Centre for Disease, P.; Control. The European Union One Health 2020 Zoonoses Report. EFSA J **2021**, 19, e06971, doi:10.2903/j.efsa.2021.6971.
4. Scheutz, F.; Teel, L.D.; Beutin, L.; Pierard, D.; Buvens, G.; Karch, H.; Mellmann, A.; Caprioli, A.; Tozzoli, R.; Morabito, S.; et al. Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. J Clin Microbiol **2012**, 50, 2951-2963, doi:10.1128/JCM.00860-12.
5. Bai, X.; Fu, S.; Zhang, J.; Fan, R.; Xu, Y.; Sun, H.; He, X.; Xu, J.; Xiong, Y. Identification and pathogenomic analysis of an Escherichia coli strain producing a novel Shiga toxin 2 subtype. Sci Rep **2018**, 8, 6756, doi:10.1038/s41598-018-25233-x.

6. Lacher, D.W.; Gangiredla, J.; Patel, I.; Elkins, C.A.; Feng, P.C. Use of the Escherichia coli Identification Microarray for Characterizing the Health Risks of Shiga Toxin-Producing Escherichia coli Isolated from Foods. J Food Prot **2016**, 79, 1656-1662, doi:10.4315/0362-028X.JFP-16-176.
7. Yang, X.; Bai, X.; Zhang, J.; Sun, H.; Fu, S.; Fan, R.; He, X.; Scheutz, F.; Matussek, A.; Xiong, Y. Escherichia coli strains producing a novel Shiga toxin 2 subtype circulate in China. Int J Med Microbiol **2020**, 310, 151377, doi:10.1016/j.ijmm.2019.151377.
8. Panel, E.B.; Koutsoumanis, K.; Allende, A.; Alvarez-Ordóñez, A.; Bover-Cid, S.; Chemaly, M.; Davies, R.; De Cesare, A.; Herman, L.; Hilbert, F.; et al. Pathogenicity assessment of Shiga toxin-producing Escherichia coli (STEC) and the public health risk posed by contamination of food with STEC. EFSA Journal **2020**, 18, e05967, doi:https://doi.org/10.2903/j.efsa.2020.5967.
9. Feldgarden, M.; Brover, V.; Gonzalez-Escalona, N.; Frye, J.G.; Haendiges, J.; Haft, D.H.; Hoffmann, M.; Pettengill, J.B.; Prasad, A.B.; Tillman, G.E.; et al. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Sci Rep **2021**, 11, 12728, doi:10.1038/s41598-021-91456-0.
10. Souvorov, A.; Agarwala, R.; Lipman, D.J. SKESA: strategic k-mer extension for scrupulous assemblies. Genome Biol **2018**, 19, 153, doi:10.1186/s13059-018-1540-z.
11. Souvorov, A.; Agarwala, R. SAUTE: sequence assembly using target enrichment. BMC Bioinformatics **2021**, 22, 375, doi:10.1186/s12859-021-04174-9.
12. Tolar, B.; Joseph, L.A.; Schroeder, M.N.; Stroika, S.; Ribot, E.M.; Hise, K.B.; Gerner-Smidt, P. An Overview of PulseNet USA Databases. Foodborne Pathog Dis **2019**, 16, 457-462, doi:10.1089/fpd.2019.2637.
13. Stevens, E.L.; Carleton, H.A.; Beal, J.; Tillman, G.E.; Lindsey, R.L.; Lauer, A.C.; Pightling, A.; Jarvis, K.G.; Ottesen, A.; Ramachandran, P.; et al. Use of Whole Genome Sequencing by the Federal Interagency Collaboration for Genomics for Food and Feed Safety in the United States. J Food Prot **2022**, 85, 755-772, doi:10.4315/JFP-21-437.
14. Poates, A.; Truong, J.; Lindsey, R.; Griswold, T.; Williams-Newkirk, A.J.; Carleton, H.; Trees, E. Sequencing of Enteric Bacteria: Library Preparation Procedure Matters for Accurate Identification and Characterization. Foodborne Pathog Dis **2022**, 19, 569-578, doi:10.1089/fpd.2022.0017.
15. Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P.A. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol **2019**, 37, 540-546, doi:10.1038/s41587-019-0072-8.
16. Wick, R.R.; Judd, L.M.; Gorrie, C.L.; Holt, K.E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol **2017**, 13, e1005595, doi:10.1371/journal.pcbi.1005595.
17. Darling, A.C.; Mau, B.; Blattner, F.R.; Perna, N.T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res **2004**, 14, 1394-1403, doi:10.1101/gr.2289704.
18. Tatusova, T.; DiCuccio, M.; Badretdin, A.; Chetvernin, V.; Nawrocki, E.P.; Zaslavsky, L.; Lomsadze, A.; Pruitt, K.D.; Borodovsky, M.; Ostell, J. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res **2016**, 44, 6614-6624, doi:10.1093/nar/gkw569.
19. Joensen, K.G.; Tetzschner, A.M.; Iguchi, A.; Aarestrup, F.M.; 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, doi:10.1128/JCM.00008-15.
20. Joensen, K.G.; Scheutz, F.; Lund, O.; Hasman, H.; Kaas, R.S.; Nielsen, E.M.; Aarestrup, F.M. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol **2014**, 52, 1501-1510, doi:10.1128/JCM.03617-13.
21. Sabol, A.; Joung, Y.J.; VanTubbergen, C.; Ale, J.; Ribot, E.M.; Trees, E. Assessment of Genetic Stability During Serial In Vitro Passage and In Vivo Carriage. Foodborne Pathog Dis **2021**, 18, 894-901, doi:10.1089/fpd.2021.0029.
22. Wirth, T.; Falush, D.; Lan, R.; Colles, F.; Mensa, P.; Wieler, L.H.; Karch, H.; Reeves, P.R.; Maiden, M.C.; Ochman, H.; et al. Sex and virulence in Escherichia coli: an evolutionary perspective. Mol Microbiol **2006**, 60, 1136-1151, doi:10.1111/j.1365-2958.2006.05172.x.
23. Hauser, J.R.; Atitkar, R.R.; Petro, C.D.; Lindsey, R.L.; Strockbine, N.; O'Brien, A.D.; Melton-Celsa, A.R. The Virulence of Escherichia coli O157:H7 Isolates in Mice Depends on Shiga Toxin Type 2a (Stx2a)-Induction and High Levels of Stx2a in Stool. Front Cell Infect Microbiol **2020**, 10, 62, doi:10.3389/fcimb.2020.00062.
24. Melton-Celsa, A.R.; Darnell, S.C.; O'Brien, A.D. Activation of Shiga-like toxins by mouse and human intestinal mucus correlates with virulence of enterohemorrhagic Escherichia coli O91:H21 isolates in orally infected, streptomycin-treated mice. Infect Immun **1996**, 64, 1569-1576, doi:10.1128/iai.64.5.1569-1576.1996.

25. Seemann, T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014, 30, 2068-2069, doi:10.1093/bioinformatics/btu153.
26. Meng, Q.; Bai, X.; Zhao, A.; Lan, R.; Du, H.; Wang, T.; Shi, C.; Yuan, X.; Bai, X.; Ji, S.; et al. Characterization of Shiga toxin-producing Escherichia coli isolated from healthy pigs in China. BMC Microbiol **2014**, 14, 5, doi:10.1186/1471-2180-14-5.
27. Malberg Tetzschner, A.M.; Johnson, J.R.; Johnston, B.D.; Lund, O.; Scheutz, F. In Silico Genotyping of Escherichia coli Isolates for Extraintestinal Virulence Genes by Use of Whole-Genome Sequencing Data. J Clin Microbiol **2020**, 58, doi:10.1128/JCM.01269-20.
28. Zhang, Y.; Liao, Y.T.; Salvador, A.; Sun, X.; Wu, V.C.H. Prediction, Diversity, and Genomic Analysis of Temperate Phages Induced From Shiga Toxin-Producing Escherichia coli Strains. Front Microbiol **2019**, 10, 3093, doi:10.3389/fmicb.2019.03093.
29. Scheutz, F.; Nielsen, E.M.; Frimodt-Moller, J.; Boisen, N.; Morabito, S.; Tozzoli, R.; Nataro, J.P.; Caprioli, A. Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing Escherichia coli O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. Euro Surveill **2011**, 16, doi:10.2807/ese.16.24.19889-en.
30. O'Brien AD, Newland JW, Miller SF, Holmes RK, Smith HW, Formal SB. Shiga-like toxin-converting phages from Escherichia coli strains that cause hemorrhagic colitis or infantile diarrhea. Science. **1984**;226(4675):694-696. doi:10.1126/science.6387911.
31. Koch C, Hertwig S, Lurz R, Appel B, Beutin L. Isolation of a lysogenic bacteriophage carrying the stx1OX3 gene, which is closely associated with Shiga toxin-producing Escherichia coli strains from sheep and humans. J. Clin. Microbiol., 2001;39(11):3992-3998. doi:10.1128/JCM.39.11.3992-3998.2001
32. Paton AW, Beutin L, Paton JC. Heterogeneity of the amino-acid sequences of Escherichia coli Shiga-like toxin type-I operons. Gene. **1995**;153(1):71-74. doi:10.1016/0378-1119(94)00777-p
33. Bürk C, Dietrich R, Açar G, Moravek M, Bülte M, Märtlbauer E. Identification and characterization of a new variant of Shiga toxin 1 in Escherichia coli ONT:H19 of bovine origin. J Clin Microbiol. **2003**;41(5):2106-2112. doi:10.1128/JCM.41.5.2106-2112.2003
34. Piérard D, Muyldermans G, Moriau L, Stevens D, Lauwers S. Identification of new verocytotoxin type 2 variant B-subunit genes in human and animal Escherichia coli isolates. J Clin Microbiol. **1998**;36(11):3317-3322. doi:10.1128/JCM.36.11.3317-3322.1998
35. Paton AW, Paton JC, Heuzenroeder MW, Goldwater PN, Manning PA. Cloning and nucleotide sequence of a variant Shiga-like toxin II gene from Escherichia coli OX3:H21 isolated from a case of sudden infant death syndrome. Microb Pathog. **1992**;13(3):225-236. doi:10.1016/0882-4010(92)90023-h
36. Persson S, Olsen KE, Ethelberg S, Scheutz F. Subtyping method for Escherichia coli shiga toxin (verocytotoxin) 2 variants and correlations to clinical manifestations. J Clin Microbiol. **2007**;45(6):2020-2024. doi:10.1128/JCM.02591-06
37. Weinstein DL, Jackson MP, Samuel JE, Holmes RK, O'Brien AD. Cloning and sequencing of a Shiga-like toxin type II variant from Escherichia coli strain responsible for edema disease of swine. J Bacteriol. **1988**;170(9):4223-4230. doi:10.1128/jb.170.9.4223-4230.1988
38. Schmidt H, Scheef J, Morabito S, Caprioli A, Wieler LH, Karch H. A new Shiga toxin 2 variant (Stx2f) from Escherichia coli isolated from pigeons. Appl Environ Microbiol. **2000**;66(3):1205-1208. doi:10.1128/AEM.66.3.1205-1208.2000
39. Leung PH, Peiris JS, Ng WW, Robins-Browne RM, Bettelheim KA, Yam WC. A newly discovered verotoxin variant, VT2g, produced by bovine verocytotoxigenic Escherichia coli. Appl Environ Microbiol. **2003**;69(12):7549-7553. doi:10.1128/AEM.69.12.7549-7553.2003
40. Meng Q, Bai X, Zhao A, Lan R, Du H, Wang T, Shi C, Yuan X, Bai X, Ji S, Jin D, Yu B, Wang Y, Sun H, Liu K, Xu J, Xiong Y. Characterization of Shiga toxin-producing Escherichia coli isolated from healthy pigs in China. BMC Microbiol. **2014** Jan 6;14:5. doi: 10.1186/1471-2180-14-5.
41. Arndt, D.,Grant, J.R.; Marcu, A.,Sajed, T.,Pon, A.,Liang, Y., Wishart, D.S. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* **2016**, *44*, W16-21, doi:10.1093/nar/gkw387.