JOURNAL OF MEDICAL MICROBIOLOGY

RESEARCH ARTICLE

Rodwell *et al.*, *Journal of Medical Microbiology* 2024;73:001790
DOI 10.1099/jmm.0.001790

MICROBIOLOGY SOCIETY

OPEN ACCESS

Clinical and public health implications of increasing notifications of LEE-negative Shiga toxin-producing *Escherichia coli* in England, 2014–2022

Ella V. Rodwell[1,2,3], David R. Greig[1], Gauri Godbole[1] and Claire Jenkins[1,3,*]

Abstract

Introduction. Shiga toxin-producing *Escherichia coli* (STEC) belong to a diverse group of gastrointestinal pathogens. The pathogenic potential of STEC is enhanced by the presence of the pathogenicity island called the Locus of Enterocyte Effacement (LEE), including the intimin encoding gene *eae*.

Gap statement. STEC serotypes O128:H2 (Clonal Complex [CC]25), O91:H14 (CC33), and O146:H21 (CC442) are consistently in the top five STEC serotypes isolated from patients reporting gastrointestinal symptoms in England. However, they are *eae*/LEE-negative and perceived to be a low risk to public health, and we know little about their microbiology and epidemiology.

Aim. We analysed clinical outcomes and genome sequencing data linked to patients infected with LEE-negative STEC belonging to CC25 (O128:H2, O21:H2), CC33 (O91:H14) and, and CC442 (O146:H21, O174:H21) in England to assess the risk to public health.

Results. There was an almost ten-fold increase between 2014 and 2022 in the detection of all STEC belonging to CC25, CC33 and CC442 (2014 *n*=38, 2022 *n*=336), and a total of 1417 cases. There was a higher proportion of female cases (55–70%) and more adults than children, with patients aged between 20–40 and >70 most at risk across the different serotypes. Symptoms were consistent across the three dominant serotypes O91:H14 (CC33), O146:H21 (CC442) and O128:H2 (CC25) (diarrhoea >75%; bloody diarrhoea 25–32%; abdominal pain 64–72%; nausea 37–45%; vomiting 10–24%; and fever 27–30%). Phylogenetic analyses revealed multiple events of acquisition and loss of different stx-encoding prophage. Additional putative virulence genes were identified including *iha*, *agn43* and *subA*.

Conclusions. Continued monitoring and surveillance of LEE-negative STEC infections is essential due to the increasing burden of infectious intestinal disease, and the risk that highly pathogenic strains may emerge following acquisition of the Shiga toxin subtypes associated with the most severe clinical outcomes.

INTRODUCTION

Shiga toxin-producing *Escherichia coli* (STEC) are a diverse group of zoonotic, foodborne gastrointestinal pathogens defined by the presence of a bacteriophage encoded Shiga toxin gene (*stx*) [1]. Analysis of the population structure of STEC has revealed that *stx* is found in a wide range of *E. coli* clonal complexes, seven gene multi-locus sequence types (MLST) and serotypes [2–4]. Certain STEC have the potential to cause haemolytic uraemic syndrome (HUS), a systemic condition characterised by microangiopathic haemolytic anaemia, thrombocytopenia, and renal failure, and can be fatal [5]. Historically, these highly pathogenic STEC have been identified by their serotype, for example, STEC O157:H7 and STEC O26:H11. However, more recently it has been established that the key factor determining pathogenic potential is the *stx* subtype [6].

Received 25 October 2023; Accepted 12 December 2023; Published 01 February 2024
Author affiliations: [1]National Infection Service, UK Health Security Agency, 61 Colindale Avenue, London, NW9 5AT, UK; [2]Warwick Medical School, University of Warwick, Coventry, CV4 7AL, UK; [3]NIHR HPRU in Gastrointestinal Infections at University of Liverpool, Liverpool, UK.
***Correspondence:** Claire Jenkins, claire.jenkins1@ukhsa.gov.uk
Keywords: Epidemiology; Genome Sequencing; non-O157 STEC; Shiga toxin-producing Escherichia coli.
Abbreviations: CC, Clonal Complex; CTSMAC, cefixime tellurite sorbitol MacConkey; ESQ, Enhanced Surveillance Questionnaire; GBRU, Gastrointestinal Bacteria Reference Unit; HUS, haemolytic uraemic syndrome; IID, infectious intestinal disease; LAA-PAI, Locus of Adhesion and Autoaggregation pathogenicity island; LEE, Locus of Enterocyte Effacement; MLST, multi-locus sequence type; NESSS, National Enhanced STEC Surveillance System; STEC, Shiga toxin-producing *Escherichia coli*; UKHSA, UK Health Security Agency.
Two supplementary tables and three supplementary figures are available with the online version of this article.

There are two types of Stx, Stx1 and Stx2, and these are further divided into subtypes, *stx1a,c-d*, *stx2a-stx2o* [6–12]. Although all the established *stx* subtypes appear to have the potential to cause gastrointestinal symptoms in humans, certain subtypes are associated with more severe clinical outcomes [6]. A higher proportion of patients infected with STEC harbouring *stx1a, stx2a* and *stx2d* report bloody diarrhoea and hospitalisation, and those patients that develop HUS are more commonly infected with STEC that have Stx subtypes *stx2a* and activatable-*stx2d* [6, 13, 14, 15]. The pathogenic potential of STEC is further enhanced by the presence of a pathogenicity island called the Locus of Enterocyte Effacement (LEE) [1, 6]. The LEE encodes a cassette of pathogenicity genes (including the intimin encoding gene *eae*) involved in attachment of *E. coli* to the gut mucosa of humans and animals, thus establishing colonisation and, in certain hosts notably humans, facilitating infection [16].

The introduction of commercial PCR for the direct detection of gastrointestinal pathogens in faecal specimens has improved diagnoses and surveillance of STEC, and detection of foodborne outbreaks [2, 17–20]. Prior to the implementation of PCR, STEC serotype O157:H7 was the only serotype of STEC routinely sought in England. The majority of STEC O157:H7 are non-sorbitol fermenting and resistant to cefixime and tellurite and can be cultured on cefixime tellurite sorbitol MacConkey (CT-SMAC) media. Most other STEC ferment sorbitol and/or are sensitive to tellurite and therefore CT-SMAC is not a suitable selective medium for non-O157 STEC. The PCR target gene for STEC is *stx*, the defining virulence factor for the STEC group and potentially all STEC serotypes, with the exception of *stx2f* in certain assays, can be detected using this approach [2, 21, 22].

Recent analysis of data held by the UK Health Security Agency (UKHSA) revealed that that the three most frequently reported LEE-negative STEC serotypes causing infectious intestinal disease (IID) in England were O128:H2, O91:H14, and O146:H21 [2, 23]. These three serotypes belonged to *E. coli* CC25, CC33 and CC442, respectively. Previous studies have shown that LEE-negative STEC belonging to the serogroup O91, including O91:H14 (CC33) and O91:H21 (CC442), were among the most common non-O157 STEC serogroups associated with human illness in Europe [24, 25]. Nüesch-Inderbinen *et al.* [26] analysed 48 clinical STEC O91 isolates collected during 2003–2019 in Switzerland and identified the Locus of Adhesion and Autoaggregation pathogenicity island (LAA-PAI) carrying *iha* and other putative pathogenicity genes, such as *hes, tspA* or *agn43* [26–28]. Other non-LAA-PAI virulence factors have been described in LEE-negative isolates, including the subtilase cytotoxin gene *subA*, colicin genes (*cba, cia, cib* and *cma*) or microcin genes (*mcmA, mchB, mchC* and *mchF*) [28–30].

Although known to cause a high proportion of STEC infection, little is known about the LEE-negative STEC in England [2, 23]. In this study, we reviewed the genome sequencing data from LEE-negative STEC belonging to CC25, CC33 and CC442 held in the UK Health Security Agency (UKHSA) archives. We linked the sequencing data to patient demographics and clinical outcomes to determine the association between virulence profile and symptoms severity, and to gain a better understanding of the populations most at risk of infection.

METHODS

Microbiology

In England, all stool samples from hospitalised patients and community acquired gastrointestinal (GI) infections are tested for STEC O157:H7 using CT-SMAC agar, and non-sorbitol fermenting colonies agglutinating with *E. coli* O157 antisera are referred to the Gastrointestinal Bacteria Reference Unit (GBRU), UKHSA for confirmation and typing. Where local diagnostic laboratories have implemented a commercial gastrointestinal PCR assay that include primers targeting *stx*, all faecal specimens are tested for *stx* to confirm the presence of STEC. *Stx*-positive faecal specimens are cultured on CTSMAC and/or CHROMagar STEC, and colonies exhibiting characteristics indicative of STEC may be referred to GBRU [31]. Alternatively, the *stx*-positive faecal specimen may be referred to GBRU for PCR and culture as described previously [21, 31]. All isolates of STEC are sequenced, including those that are referred to GBRU from local diagnostic laboratories, and those that are cultured at GBRU from referred faecal specimens.

DNA extraction and genomic processing

Genomic DNA was extracted and sequenced on Illumina HiSeq 2500 and NextSeq 1000 platforms at UKHSA and held in the in-house data repository. Post whole genome sequencing (WGS), data are processed through an in-house pipeline that determines serotype using GeneFinder (https://github.com/phe-bioinformatics/gene_finder) [32]. The Stx subtype was determined by aligning putative reads to *stx* reference genes and detection of unique *stx*-subtype SNP positions [33]. Multilocus sequence typing (MLST) was performed using Metric Orientated Sequence Typer (https://github.com/phe-bioinformatics/MOST) [34].

Additional *in silico* gene profiling of isolates was performed to investigate the presence of putative virulence factors and genes associated with the LAA. The region 385984–472336 bp of *Escherichia coli* O91:H21 str. B2F1 strain (contig 25) (AFDQ01000026.1) was downloaded from the NCBI nucelotide database [27, 35]. Bakta (v1.6.1) was used to annotate this genome region and the genes from the .ffn file were made into a custom database following previous studies [27]. Virulence genes previously described [26–30] were extracted from the Centre for Genomic Epidemiology VirulenceFinder *E. coli* database (v.2022.12.02) and curated into a custom database for genefinder. Read mapping was conducted using Illumina short reads, depth coverage was set to 5% and homology/coverage values set to >85 to indicate detection.

SnapperDB v0.2.8 is the UKHSA in-house database that holds called variant data, achieved from genomic DNA sequencing, relative to an appropriate reference for *E. coli* CCs [36]. SnapperDB v0.2.6 was employed to generate a whole genome alignment of isolates representing the t:250 level of CC33 and CC442, and strain level for CC25, based on a variant alignment to reference strains *E. coli* O91:H14 (SRR3581435), *E. coli* O146:H21 (JASR01000001.1) and *E. coli* O128:H2 (SRR3578969) respectively. Gubbins v2.0.0 [37] was used on the whole genome alignment to identify recombinant regions, which were then masked during the building of a second alignment of genomes within this study, where a variant position belongs to a minimum of 80% of strains in the alignment. This alignment was examined by IQTree v2.0.4 [38] under the best-fit model which produced a maximum-likelihood phylogeny that was visualised in ITOL v5.7 [39].

Epidemiological data collection

The UKHSA STEC Operational guidance [40] recommends that the STEC Enhanced Surveillance Questionnaires (ESQ) is administered to all cases, however, cases of STEC O157, and patients infected with STEC harbouring *stx2* and *eae* or *aggR*, and/or children under the age of six are prioritised. The collected data includes travel and food histories, environmental and animal exposures and clinical symptoms. These epidemiological data are stored in the National Enhanced STEC Surveillance System (NESSS) is paired with genomic and microbiological data for each case. Patient data, including age, sex, clinical presentation, and travel history, was reviewed. Cases are defined as an individual patient and an isolate is defined as a strain originating from a case, whereby a case can have multiple isolates. Multiple isolates can arise where a strain is unique at the t5 level or with differing *stx* profiles, per case.

Data deposition

All isolates reviewed in this study are deposited under Bioproject PRJNA315192 as part of routine surveillance at UKHSA. Strain metadata is found in Table S1, available in the online version of this article.

RESULTS

Clonal complex microbiology and typing data

There were 1451 isolates belonging to either CC25 (*n*=323), CC33 (*n*=538) or CC442 (*n*=590) submitted to GBRU between 2014 and 2022 from 1417 individual cases (CC25 *n*=315; CC33 *n*=528; CC442 *n*=574). With the exception of CC33 in 2020, there has been a year-on-year increase in notifications of each CC, with nearly ten times more notifications in 2022 (*n*=336 in total; CC25 *n*=81; CC33 *n*=122; CC442 *n*=133) than in 2014 (*n*=38 in total; CC25 *n*=8; CC33 *n*=14; CC442 *n*=16) (Fig. 1).

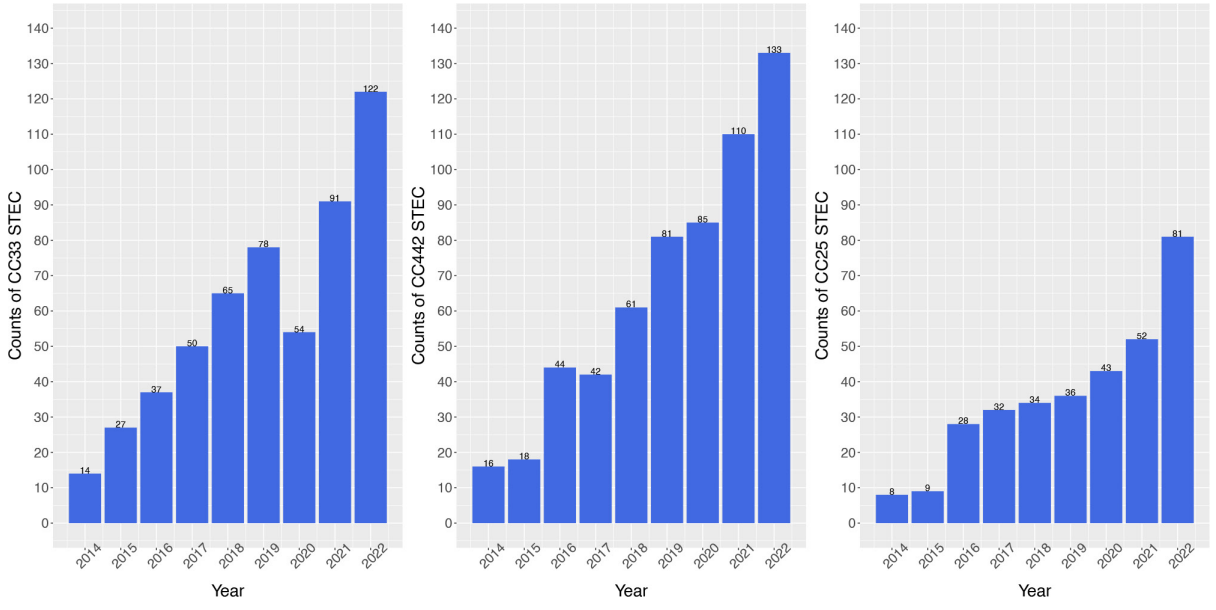

Fig. 1. Notifications of STEC infections belonging to CC33 (*n*=538), CC442 (*n*=590) and CC25 (*n*=323). Number of notifications is represented using cumulative counts of unique isolates each year from 2014 to 2022 from the patient sample date. An isolate is unique at the t5 level or with differing *stx* subtype per case. CC33 represented as the left figure, CC442 in the middle and CC25 as the right.

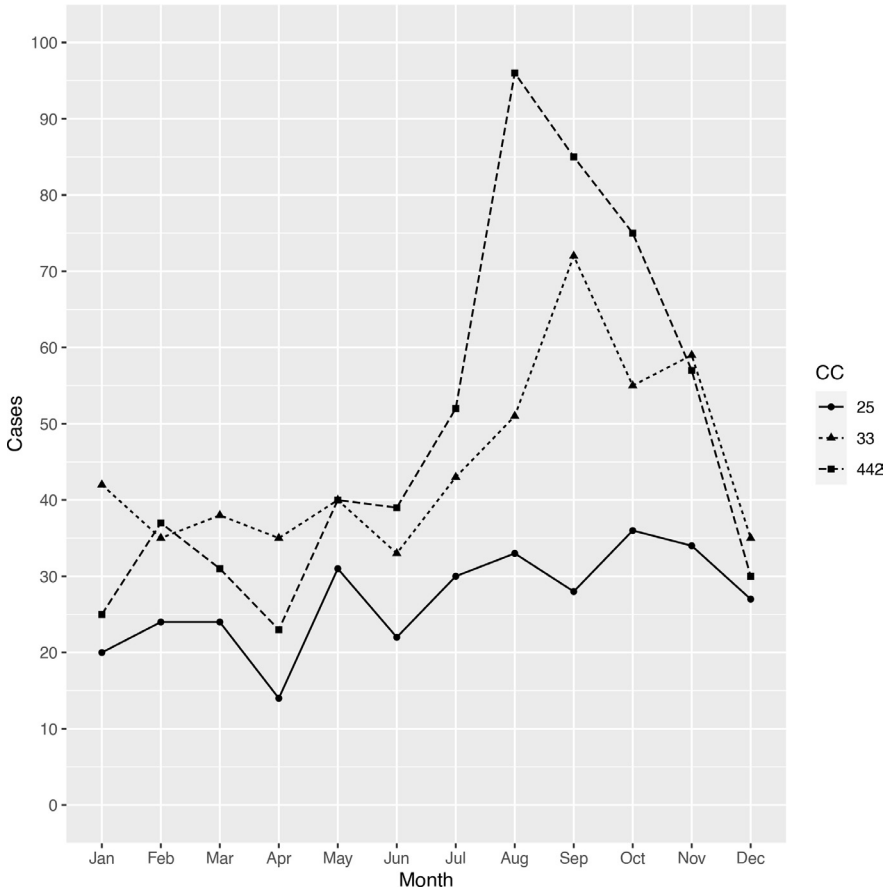

Fig. 2. Seasonality of STEC infections belonging to CC33 (*n*=538), CC442 (*n*=590) and CC25 (*n*=323). Seasonality is represented using cumulative counts of unique isolates per month from 2014 to 2022, from the patient sample date. An isolate is unique at the t5 level or with differing *stx* subtype, per case.

All three CCs exhibit an increase in case numbers during late summer, however the rise is steeper for CC33 and CC442 (Fig. 2). As observed with other non-O157 STEC CCs in England [17–19], London (CC25 *n*=103/315, 33%, CC33 *n*=189/528, 36%, CC442 *n*=184/574, 32%) and the South East of England (CC442 *n*=200/574, 35%; CC33 *n*=180/528, 34%, CC25 *n*=118/315, 37%) were the two regions reporting the highest numbers of cases of LEE-negative STEC (Fig. 3).

Although each clonal complex comprised more than one serotype, each were characterised by a dominant type. The dominant STEC serotype in CC25 was O128:H2 (*n*=291/323, 90%), followed by O21:H2 (24/323, 7%), O unidentifiable (7/323, 2%) and O8:H2 (1/323, 0.3%). Serotype O91:H14 was most frequently detected in CC33 (*n*=518/538, 96%), followed by O unidentifiable:H14 (15/538), O117:H14 (4/538) and O91:H unidentifiable (1/538). In CC442, serotype O146:H21 (*n*=553/590, 94%) was most common, followed by O174:H21 (18/590), O unidentifiable:H21 (16/590), O91:H21 (2/590) and O156:H21 (1/590) (Figs 4–6). The isolates within CC33 and CC442 almost exclusively belonged to ST33 (537/538, 99.8%) and ST442 (570/590, 97%) respectively, whereas in CC25, ST25 (239/323, 74%) and ST4748 (60/323 19%) were the most frequently detected sequence types (Figs 4–6).

Within CC33, the majority of isolates had the *stx* subtypes *stx1a* and/or *stx2b* either together (*stx1a, stx2b n*=404/538, 75%), or alone (*stx1a* only *n*=33/538, 6%; *stx2b* only *n*=94/538, 17%) (Fig. 5). The most common *stx* profiles detected in CC442 were *stx1c* (*n*=271/590, 46%), *stx1c, stx2b* (*n*=224/590, 38%) or *stx2b* (*n*=56/590, 9%) (Fig. 6). Like CC442, most isolates in CC25 had *stx1c* (*n*=25/323, 8%), *stx2b* (*n*=85/323, 26%) or both *stx1c, stx2b* (*n*=210/323, 65%) (Figs 4–6).

Epidemiology of LEE-negative STEC serotypes

To determine serotype specific epidemiological signals within these complexes, data was extracted for O128:H2 and O21:H2 (CC25), O91:H14 (CC33), and O146:H21 and O174:H21 (CC442). All key serotypes associated with all three CCs had a higher proportion of female cases (O128:H2 *n*=174/283, 61%; O21:H2 *n*=16/23, 70%; O91:H14 *n*=300/506, 59%; O146:H21 *n*=295/540,

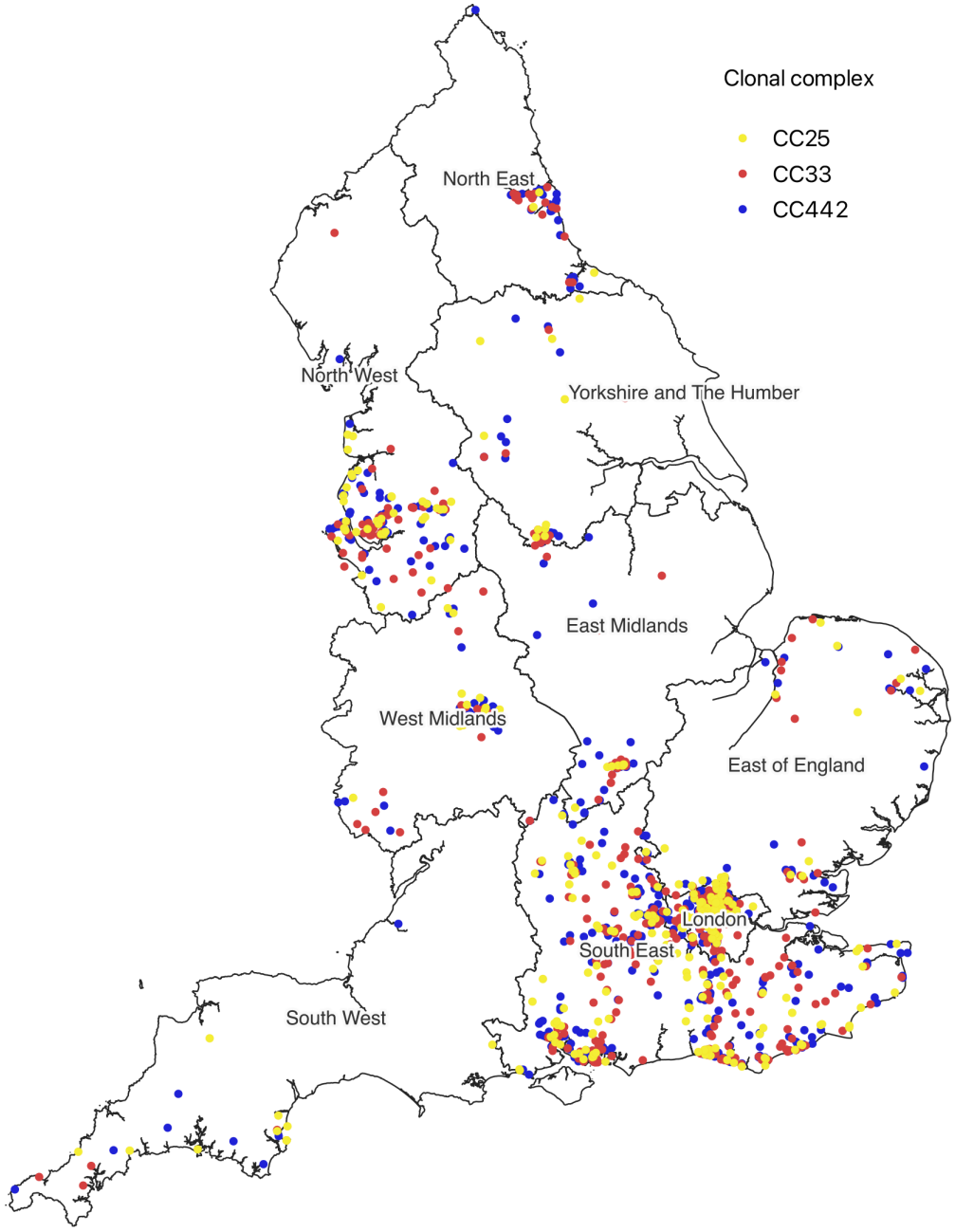

Fig. 3. Geographical distribution of cases of CC442, CC33 and CC25 STEC cases in England based on patient postcodes. Where postcode was not available, no data for that case was added.

55%, O174:H21 *n*=12/18, 67%) and adults were more often infected than children, with patients aged 20–40 and over 70 years old seemingly most at risk (Table 1 and Fig. 7).

Of the 1374 cases infected with serotypes O146:H21, O91:H14, O128:H2, O174:H21 and O21:H2, data on clinical outcomes were available for 463/1374 (34%) symptomatic cases (O128:H2 *n*=92/285, 33%; O21:H2 *n*=10/23, 43%; O91:H14 *n*=168/508, 33%; O146:H21 *n*=184/540, 34%; O174:H21 *n*=9/18, 50%) (Table 2). Diarrhoea was reported by over 75% of cases in all three CCs regardless of serotype. The proportion of cases reporting bloody diarrhoea (25–32%), abdominal pain (64–72%), nausea (37–45%), vomiting (10–24%) and fever (27–30%) were also consistent across the three main serotypes (O146:H21, O91:H14, O128:H2). Serotype O174:H21 (CC442) presented more atypical with no bloody diarrhoea, reduced abdominal pain (44%). Additionally, there were two cases of HUS (2/9, 22%) associated with STEC O174:H21; both were males (aged 5 months and 62 years old) harbouring *stx2c* with or without *stx2d*. Although the literature has demonstrated the pathogenicity of STEC O91:H21

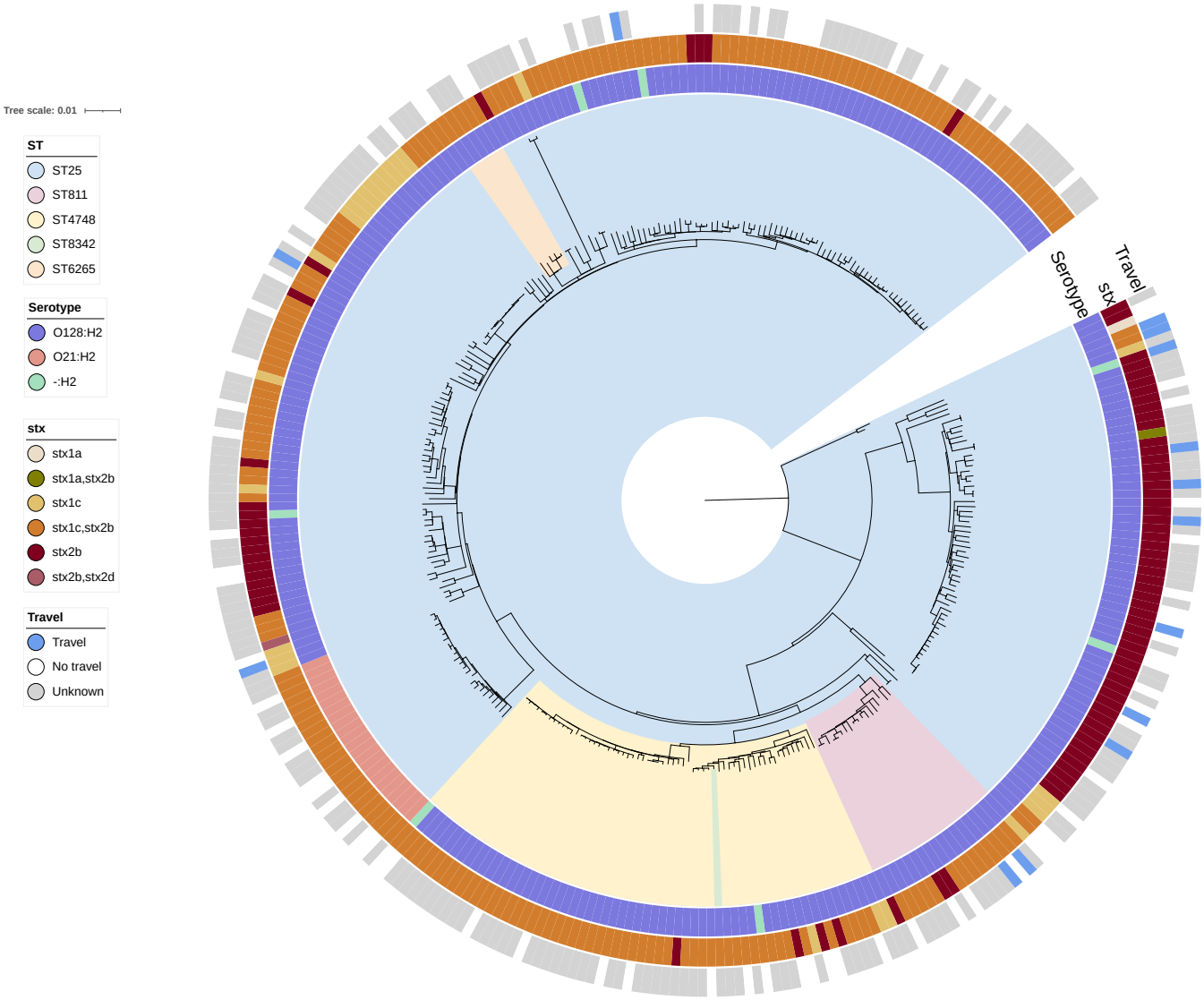

Fig. 4. Clonal Complex 25. Phylogenetic analysis of STEC isolates belonging to CC25 in England. Figure was generated using Snapper DB, IQTree2 and visualised in ITOL. The phylogeny Is rooted at midpoint and the alignment was length 26074 bp. Annotations are ST, serotype, *stx* subtype and travel association (where available).

[24], in this study STEC O91:H21 (*n*=2) was not associated with bloody diarrhoea, hospitalisation or HUS, despite the isolates harbouring *stx2a* and *stx2d*.

Travel data

Of the CC25 cases with travel data (*n*=116), there were 14 cases (12%) reporting recent travel outside of the UK, with destinations where stated included Spain (*n*=3), France (*n*=2), Italy (*n*=2) or Turkey (*n*=2). There were 14/183 (8%) cases of CC33 reporting recent travel outside the UK to Europe (*n*=10), Asia (*n*=1), North America (*n*=2) and Oceania (*n*=1), and 15/199 cases of CC442 (8%) determined to be travel associated (Europe *n*=12; North Africa *n*=3). Isolates from the majority of travel-associated cases of CC25 (*n*=9/14, 64%) were located within a clade within ST25 characterised by the presence of *stx2b* (Fig. 4), whereas cases of CC33 and CC442 with confirmed travel were dispersed across the phylogeny (Figs 5 and 6).

Phylogenetic analysis and virulence of isolates

To gain context on the relatedness and genomic differences of isolates in this study, per-CC phylogenetic analyses were performed. The population structure of CC25, CC33 and CC442 revealed multiple events of acquisition and loss of *stx*-encoding prophage across the phylogenies. CC25 comprised five different STs, although 90% of isolates belonged to serotype O128:H2 despite variation in ST

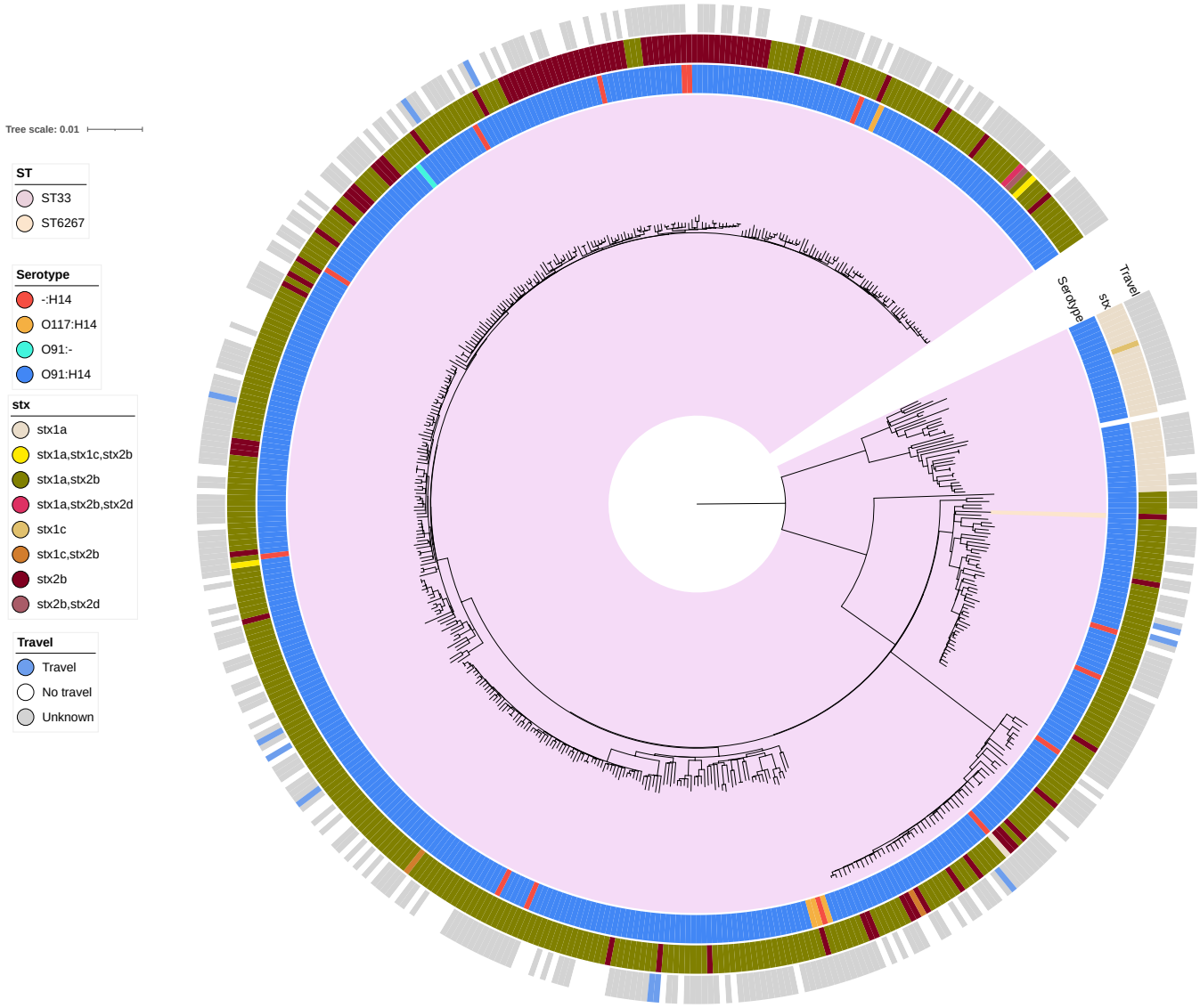

Fig. 5. Clonal Complex 33. Phylogenetic analysis of STEC isolates belonging to CC33 in England. Figure was generated using Snapper DB, IQTree2 and visualised in ITOL. The phylogeny Is rooted at midpoint and the alignment was length 30614 bp. Annotations are ST, serotype, *stx* subtype and travel association (where available).

(Fig. 4). Isolates of STEC O21:H2 belonged to ST25, which was also the ST of the majority of isolates of serotype O128:H2. Most isolates within CC25 had *stx1c, stx2b* or both (Fig. 4). CC33 comprised two main clades, one characterised by the presence of *stx1a* and the other by mainly *stx1a, stx2b*, or *stx2b* only (Fig. 5). There was evidence of limited, sporadic O or H antigen gene switching, but for the most part the isolates belonged to serotype O91:H14 (Fig. 5). The two main clades of CC442 were characterised by different serotypes, O174:H21 and O146:H21 (Fig. 6). The *stx* subtypes across CC442 were variable but for STEC O174:H21 most had *stx2c*, with or without *stx2d*, and for STEC O146:H21, like CC25, most had *stx1c*, *stx2b* or both (Fig. 6).

In this study, gene presence is described within each CC to identify key differences between CC25, CC33 and CC442. Further serotype specific and enhanced gene detection data is listed in Table S2. Previous studies have shown that *iha* is found on a pathogenicity island called the LAA [27, 28, 30]. The complete LAA comprised four modules and *iha* is found on module 2, and genes found on other modules include *hes*, *tpsA* and *agn43* [27, 28, 30]. Most isolates across all three CCs had *iha* (CC25 316/320, 99%; CC33 533/536, 99%; CC442 586/586, 100%) and *agn43* (CC25 313/320, 98%; CC33 460/536, 86%; CC442 379/586, 65%) whereas the presence *tpsA* varied depending on CC and was most prevalent in CC442 (CC25 60/320, 19%; CC33 20/536, 4%; CC442 500/586, 85%) and *hes* was rarely detected (CC25 2/320, 1%; CC33 32/536, 6%; CC442 22/586, 4%) (Table S2 and Figs S1, S2 & S3).

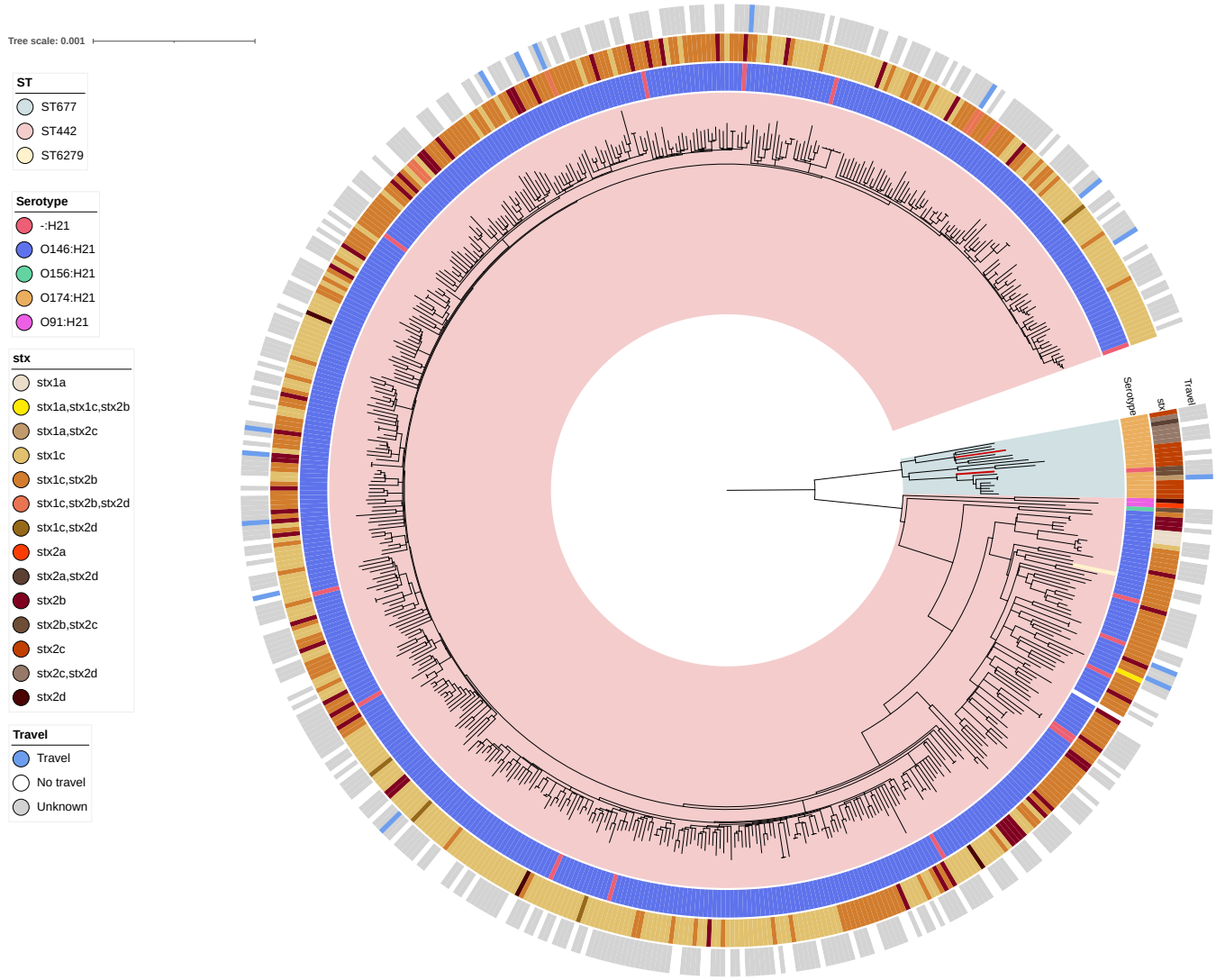

Fig. 6. Clonal Complex 442. Phylogenetic analysis of STEC isolates belonging to CC442 in England. The two STEC O174:H21 isolated from patients diagnosed with HUS are indicated with a red branch. Figure was generated using Snapper DB, IQTree2 and visualised in ITOL. The phylogeny Is rooted at midpoint and the alignment was length 14837 bp. Annotations are ST, serotype, *stx* subtype and travel association (where available).

The subtilase toxin gene *subA* was present across CC33, CC442 and CC25, (505/536, 94%; 560/586, 96%; 318/320, 99% respectively), whereas the enterohaemolysin gene *ehxA* was most commonly found in CC442 (567/586, 97%) and CC25 (264/320, 83%) compared to CC33 (301/536, 56%) (Table 2, Figs S1, S2 & S3). Although *subA* and *ehxA* were detected in most isolates within CC442, neither of these genes were detected in the two STEC O174:H21 isolated from the two cases diagnosed with HUS (Table S2 and Figs S1, S2 & S3). Of the three CCs described here only CC442 consistently had genes (*terB, terC, terD, terE, terF*) encoding resistance to tellurite (526/586, 90%) compared to CC25 (8/320, 3%) and CC33 (0/536, 0%).

DISCUSSION

In England, STEC surveillance strategies focus on the detection of cases infected with STEC that have enhanced virulence, specifically the presence of *stx2a* or *stx2d* and *eae*, a marker for the LEE locus [40]. This is to ensure that where resources are limited, clinical and public health investigations prioritise patients reporting symptoms at the more severe end of the clinical spectrum [41]. However, all faecal specimens that are PCR positive for STEC and referred to UKHSA, are cultured for all STEC serotypes [21, 31]. Previous studies and routine surveillance activities have identified LEE-negative STEC belonging to CC25, CC33 and CC442, as causing a high proportion of STEC infections in England [2, 23].

Table 1. Age and sex of cases of key serotypes of CC33: STEC O91:H14 (*n*=506), CC25: STEC O128:H2 (*n*=283) and STEC O21:H2 (*n*=23) and CC442: STEC O146:H21 (*n*=540) and STEC O174:H21 (*n*=18). Age (was determined from sample date or receipt date) and sex of the patient (where both were available) was used on a per case basis

	Clonal complex 33			Clonal complex 442						Clonal complex 25							
	STEC O91:H14			STEC O146:H21			STEC O174:H21			STEC O128:H2			STEC O21:H2			Total	
	M	F	Total	M	F	Total	M	F	Total	M	F	Total	M	F	Total	Total	%
70 +	32	47	**79**	46	55	**101**	1	4	**5**	13	38	**51**	2	1	**3**	**239**	**17**
60–69	16	27	**43**	26	26	**52**	1	1	**2**	14	12	**26**	1	3	**4**	**127**	**9**
50–59	28	39	**67**	36	32	**68**	2	0	**2**	14	10	**24**	3	1	**4**	**165**	**12**
40–49	25	38	**63**	23	27	**50**	0	2	**2**	15	24	**39**	0	2	**2**	**156**	**11**
30–39	32	48	**80**	37	46	**83**	0	4	**4**	12	35	**47**	0	3	**3**	**217**	**16**
20–29	19	60	**79**	27	45	**72**	0	1	**1**	14	30	**44**	0	6	**6**	**202**	**15**
10–19	30	27	**57**	24	33	**57**	1	0	**1**	8	10	**18**	1	0	**1**	**134**	**10**
0–9	24	14	**38**	26	31	**57**	1	0	**1**	19	15	**34**	0	0	**0**	**130**	**9**
Total	**206**	**300**	**506**	**245**	**295**	**540**	**6**	**12**	**18**	**109**	**174**	**283**	**7**	**16**	**23**	**1370**	**100**

Like STEC O157:H7 (CC11) and STEC O26:H11 (CC29), STEC belonging to CC25, CC33 and CC442 are prevalent in ruminants and transmitted to humans via direct contact with animals or consumption of contaminated food. Strains of STEC O91:H14 (ST33) are frequently isolated from cattle (dairy and beef) [28] and sheep [42], and have been isolated from retail meats [25, 28] and raw pet food [43]. STEC O146:H21 and O91:H21 (ST442) have been found in a variety of animal reservoirs including roe deer [29, 44, 45], fox, wild boar [44, 45], sheep [46], and in raw pet food [43] and shellfish [47]. STEC O128:H2 (CC25) have been isolated from sheep, deer and calves [48–51].

Unlike other STEC CCs we have reviewed [19, 52, 53], there were no travel associated clades. However, it should be noted that travel data was poorly captured because of the lack of public health follow-up of cases infected with LEE-negative STEC. Where travel histories were available, most cases were associated with domestic acquisition, suggesting that STEC O128:H2, O91:H14 and O146:H21 are endemic in the ruminant populations in the UK.

Like other STEC CCs in England [17–19, 54], we identified a higher proportion of females cases of CC25, CC33 and CC442 than male (~55–70%), and historically this has been attributed to the increased risk of exposure to STEC during food preparation and childcare, although these are no longer regarded as traditionally female roles. Unlike other STEC CCs we have studied, more notifications of CC25, CC33 and CC442 occur in adults than children [17–19, 54]. We cannot be certain why but we can speculate that it may be due to different environmental exposures, different diet, waning immunity and/or changes in the gut microbiome with age.

In common with other non-O157 STEC clonal complexes in England [17–19], notifications of STEC belonging to CC25, CC33 and CC442 have increased every year since 2014, most likely due to the increasing number of frontline diagnostic laboratories implementing commercial PCR assays that target *stx* [2]. However, it is possible that this change in the diagnostic algorithm is masking a true increase in non-O157 STEC in England. We noted a decrease in notifications of CC33 in 2020 that may have been the result of social distancing measures during the COVID-19 pandemic, however, this decrease was not observed in notifications of STEC belonging to CC25 or CC442. This may reflect differences in the source and/or transmission route of CC33, compared to CC25 and CC442, although it is unclear what these differences might be. Travel abroad was restricted during the pandemic but based on available data the proportion of cases reporting travel were similar for CC33 and CC442. Previous studies have shown that CC442 may be associated with colonisation of a broad range of animal hosts and may be ubiquitous in the environment [29, 43–46, 55]. Survival in the environment and persistence in the animal reservoir and the food chain have been linked to the presence of genes conferring resistance to tellurite [56, 57]. However, it is unclear whether these factors contributed to the ability of isolates belonging to CC442 to continue to cause infections in humans during the pandemic.

The LEE-negative CCs described in this study exhibited a similar seasonal pattern to other STEC serotypes we have reviewed [17–19, 54]. The increase in notifications of CC442 and CC33 in late summer/early autumn is typical of STEC. CC25 exhibits less pronounced seasonality and, again, this may reflect bespoke source and/or transmission routes specific to this type. Other studies have reported an association between isolates belonging to CC25 and prevalence in the ovine reservoir, whereas CC442 has a wider host range [29, 43–46, 48–51, 55]. However, even if this association was confirmed, it is unclear how this might impact on seasonality patterns.

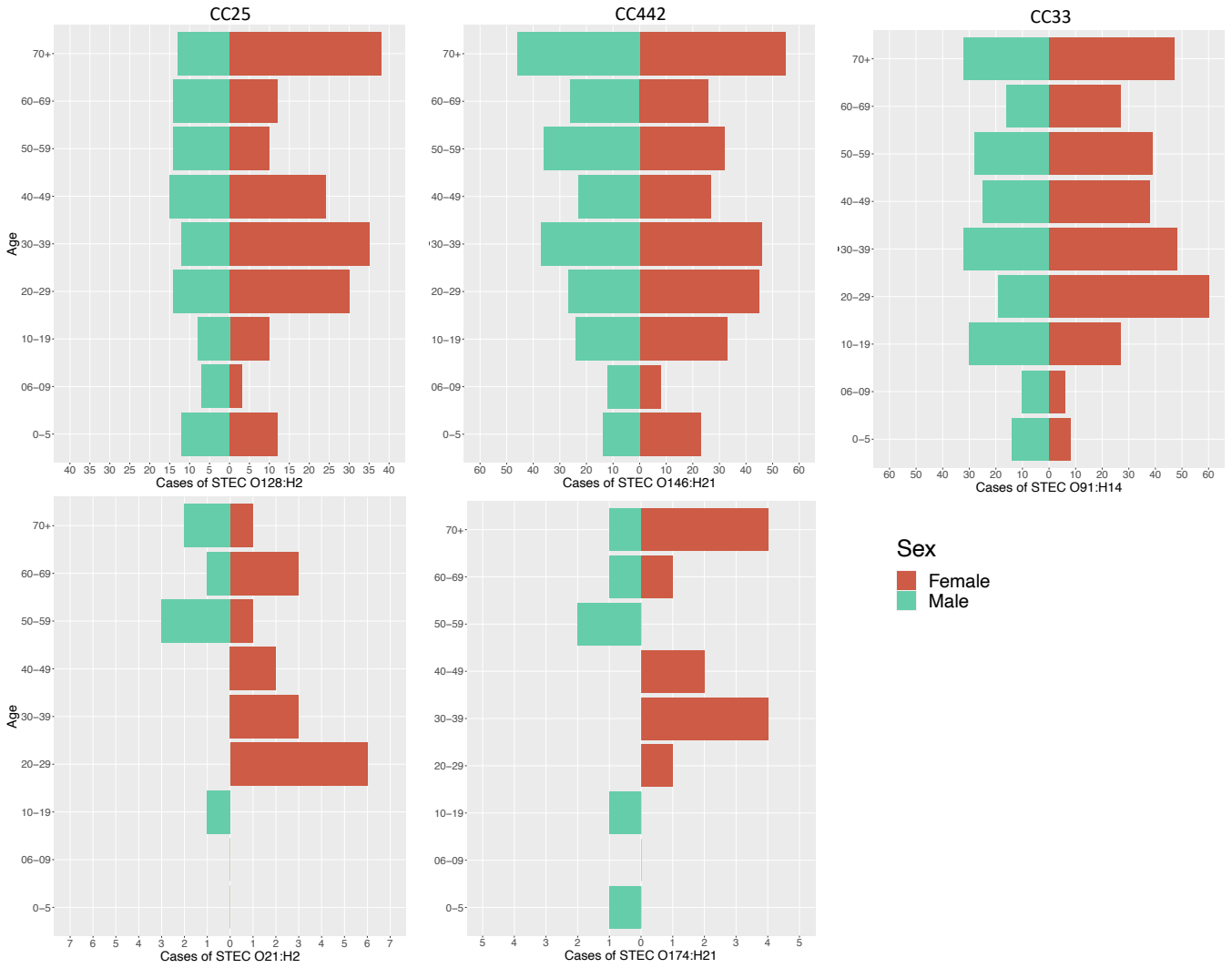

Fig. 7. Age-sex of cases infected with key serotypes from the three CCs. CC32: STEC O128:H2 (*n*=283) and STEC O21:H2 (*n*=23), CC442: STEC O146:H21 (*n*=540) and STEC O174:H2 (*n*=18), and CC33: STEC O91:H14 (*n*=506). Age (was determined from sample date or receipt date) and sex of the patient (where both were available) was used on a per case basis. Age groups read bottom to top as 0–5, 6–9, 10–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70+. Serotypes are labelled under the clonal complexes vertically. Male indicated as green on the left of the pyramid and females indicated as red on the right of the pyramid.

The population structure of all three CC reviewed in this study showed multiple examples of acquisition and loss of a variety of *stx* subtypes across each phylogeny. Most isolates belonging to CC33 had *stx1a* and/or *stx2b*, and most isolates belonging to CC25 and CC442 had *stx1c* and/or *stx2b*. Previous studies have shown that patients infected with LEE-negative STEC are less likely to develop HUS that those infected with LEE-positive STEC [6]. In our dataset of 463 cases, only two were diagnosed with HUS, both belonged to O174:H21 within CC442. This serotype has been previously described as having enhanced pathogenic potential, perhaps because of the associated *stx* subtypes [58, 59]. Unlike most isolates in CC442, STEC O174:H21 had *stx2c* and/or *stx2d*, both subtypes that have been associated with progression to HUS. Other putative virulence factors that were consistently detected across all three CCs were the LAA encoded genes *iha* and *agn43,* thought to be involved in adherence of the bacteria to the host gut mucosa, and the subtilase toxin *subA* [26–28, 30]. Both of the STEC O174:H21 isolated from the two HUS cases had *iha* but not *subA* or *ehxA*. The putative adherence factors located on the LAA, *hes*, *tpsA*, and the haemolysin *ehxA*, were present in certain CCs and serotypes but not in others, demonstrating the modular appearance of the LAA as described by Montero *et al.* [27].

Over the last decade, UKHSA surveillance systems have detected a ten-fold increase in notifications of STEC belonging to CC25, CC33 and CC442. These LEE-negative strains exhibit the same seasonality patterns and share the same animal reservoirs and transmission routes as the LEE-positive STEC CCs we have reviewed to date [17–19, 54]. However, unlike the LEE-positive STEC, the LEE-negative strains are more commonly detected in adults than children and are rarely detected in cases of HUS, possibly because most isolates do

Table 2. Clinical presentation of symptomatic cases of CC33: STEC O91:H14 (*n*=168), CC442: STEC O146:H21 (*n*=184) and STEC O174:H14 (*n*=9), CC25: STEC O128:H2 (*n*=92) and STEC O21:H2 (*n*=10). Table indicates where patient self-reported clinical symptom or outcome in the ESQ

	Clonal complex 33		**Clonal complex 442**				**Clonal complex 25**			
	O91:H14 (*n*=168)		O146:H21 (*n*=184)		O174:H21 (*n*=9)		O128:H2 (*n*=92)		O21:H2 (*n*=10)	
Clinical Presentation	Yes	%	Yes	%	Yes	%	Yes	%	Yes	%
Diarrhoea	127	76	155	84	8	89	77	84	8	80
Abdominal pain	121	72	119	64	4	44	64	70	5	50
Nausea	75	45	68	37	4	44	37	40	2	20
Fever	46	27	49	27	4	44	28	30	2	20
Vomiting	40	24	42	23	2	22	9	10	2	20
Blood in stool	45	27	46	25	0	0	29	32	3	30
Admitted to hospital	20	12	22	12	3	33	12	13	1	10
HUS	0	0	0	0	2	22	0	0	0	0
Died	0	0	0	0	0	0	0	0	0	0

not carry the *stx* subtypes known to be associated with progression to HUS [60]. Although STEC CC25, CC33 and CC442 lack the LEE pathogenicity island, they may have alternative mechanisms for adhering to the host gut mucosa [26–28, 30]. We identified two cases of HUS linked to a small clade within CC442, characterised by the acquisition of *stx2c* and/or *stx2d*-encoding bacteriophage, and it appears that this event may have been sufficient to enhance the pathogenic potential of the LEE-negative strains. The follow up data collection rate for STEC without the virulence profile *eae*$^{+}$/*stx2*$^{+}$ is low as demonstrated in this study (approximately 34%), meaning a potential bias could be associated with the clinical presentation data. However, highly pathogenic strains may emerge following acquisition of the Shiga toxin subtypes associated with the most severe clinical outcomes. Therefore, continued monitoring and surveillance of LEE-negative STEC infections is essential to fully uncover the true burden of infectious intestinal disease caused by these zoonotic, foodborne pathogens.

Funding information
This study is funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Gastrointestinal Infections, a partnership between the UK Health Security Agency, the University of Liverpool and the University of Warwick. The views expressed are those of the author(s) and not necessarily those of the NIHR, the UK Health Security Agency or the Department of Health and Social Care.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
The authors declare that there is no requirement for ethical approval for this submission. This work was undertaken to inform the delivery of patient care and to prevent the spread of infection, defined as USUAL PRACTICE in public health and health protection.

References

1. Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, *et al*. Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clin Microbiol Rev* 2013;26:822–880.
2. Vishram B, Jenkins C, Greig DR, Godbole G, Carroll K, *et al*. The emerging importance of Shiga toxin-producing *Escherichia coli* other than serogroup O157 in England. *J Med Microbiol* 2021;70:001375.
3. Gentle A, Day MR, Hopkins KL, Godbole G, Jenkins C. Antimicrobial resistance in Shiga toxin-producing *Escherichia coli* other than serotype O157:H7 in England, 2014–2016. *J Med Microbiol* 2020;69:379–386.
4. Wirth T, Falush D, Lan R, Colles F, Mensa P, *et al*. Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol Microbiol* 2006;60:1136–1151.
5. Jenkins C, Byrne L, Vishram B, Sawyer C, Balasegaram S, *et al*. Shiga toxin-producing *Escherichia coli* haemolytic uraemic syndrome (STEC-HUS): diagnosis, surveillance and public-health management in England. *J Med Microbiol* 2020;69:1034–1036.
6. Koutsoumanis K, Allende A, Alvarez-Ordóñez A, Bover-Cid S, Chemaly M, *et al*. Pathogenicity assessment of Shiga toxin-producing *Escherichia coli* (STEC) and the public health risk posed by contamination of food with STEC. *EFS2* 2020;18.
7. Scheutz F, Teel LD, Beutin L, Piérard D, Buvens G, *et al*. Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J Clin Microbiol* 2012;50:2951–2963.
8. Bai X, Fu S, Zhang J, Fan R, Xu Y, *et al*. Identification and pathogenomic analysis of an *Escherichia coli* strain producing a novel Shiga toxin 2 subtype. *Sci Rep* 2018;8:6756.
9. Yang X, Bai X, Zhang J, Sun H, Fu S, *et al*. *Escherichia coli* strains producing a novel Shiga toxin 2 subtype circulate in China. *Int J Med Microbiol* 2020;310:151377.
10. Bai X, Scheutz F, Dahlgren HM, Hedenström I, Jernberg C. Characterization of clinical *Escherichia coli* strains producing a novel shiga toxin 2 subtype in Sweden and Denmark. *Microorganisms* 2021;9:2374.

11. Gill A, Dussault F, McMahon T, Petronella N, Wang X, *et al*. Characterization of atypical shiga toxin gene sequences and description of Stx2j, a new subtype. *J Clin Microbiol* 2022;60:e0222921.
12. Lindsey RL, Prasad A, Feldgarden M, Gonzalez-Escalona N, Kapsak C, *et al*. 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. *Microorganisms*2023;11.
13. Byrne L, Adams N, Jenkins C. Association between shiga toxin-producing *Escherichia coli* O157:H7 stx gene subtype and disease severity, England, 2009-2019. *Emerg Infect Dis* 2020;26:2394–2400.
14. Melton-Celsa AR, Darnell SC, O'Brien AD. 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.
15. Bielaszewska M, Friedrich AW, Aldick T, Schürk-Bulgrin 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.
16. Elliott SJ, Sperandio V, Girón JA, Shin S, Mellies JL, *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.
17. Rodwell EV, Vishram B, Smith R, Browning L, Smith-Palmer A, *et al*. Epidemiology and genomic analysis of shiga toxin-producing *Escherichia coli* clonal complex 165 in the UK. *J Med Microbiol* 2021;70:001471.
18. Rodwell EV, Chan Y-W, Sawyer C, Carroll A, McNamara E, *et al*. Shiga toxin-producing *Escherichia coli* clonal complex 32, including serotype O145:H28, in the UK and Ireland. *J Med Microbiol* 2022;71.
19. Rodwell EV, Simpson A, Chan Y-W, Godbole G, McCarthy ND, *et al*. The epidemiology of shiga toxin-producing *Escherichia coli* O26:H11 (clonal complex 29) in England, 2014-2021. *J Infect* 2023;86:552–562.
20. Butt S, Allison L, Vishram B, Greig DR, Aird H, *et al*. Epidemiological investigations identified an outbreak of shiga toxin-producing *Escherichia coli* serotype O26:H11 associated with pre-packed sandwiches. *Epidemiol Infect* 2021;149:e178.
21. Jenkins C, Lawson AJ, Cheasty T, Willshaw GA. Assessment of a real-time PCR for the detection and characterization of verocytotoxigenic *Escherichia coli*. *J Med Microbiol* 2012;61:1082–1085.
22. Feng PCH, Jinneman K, Scheutz F, Monday SR. Specificity of PCR and serological assays in the detection of *Escherichia coli* shiga toxin subtypes. *Appl Environ Microbiol* 2011;77:6699–6702.
23. UKHSA. Shiga toxin-producing *Escherichia coli* (STEC) data: 2020; 2023. https://www.gov.uk/government/publications/escherichia-coli-e-coli-o157-annual-totals/shiga-toxin-producing-escherichia-coli-stec-data-2020
24. Mellmann A, Fruth A, Friedrich AW, Wieler LH, Harmsen D, *et al*. Phylogeny and disease association of shiga toxin-producing *Escherichia coli* O91. *Emerg Infect Dis* 2009;15:1474–1477.
25. Lee JB, Kim S-K, Wi SM, Cho Y-J, Hahn T-W, *et al*. Molecular epidemiology of sequence type 33 of shiga toxin-producing *Escherichia coli* O91:H14 isolates from human patients and retail meats in Korea. *J Vet Sci* 2019;20:87–90.
26. Nüesch-Inderbinen M, Stevens MJA, Cernela N, Müller A, Biggel M, *et al*. Distribution of virulence factors, antimicrobial resistance genes and phylogenetic relatedness among Shiga toxin-producing *Escherichia coli* serogroup O91 from human infections. *Int J Med Microbiol* 2021;311:151541.
27. Montero DA, Velasco J, Del Canto F, Puente JL, Padola NL, *et al*. Locus of Adhesion and Autoaggregation (LAA), a pathogenicity island present in emerging shiga toxin-producing *Escherichia coli* strains. *Sci Rep* 2017;7:7011.
28. Colello R, Vélez MV, González J, Montero DA, Bustamante AV, *et al*. First report of the distribution of Locus of Adhesion and Autoaggregation (LAA) pathogenicity island in LEE-negative Shiga toxin-producing *Escherichia coli* isolates from Argentina. *Microb Pathog* 2018;123:259–263.
29. Dias D, Costa S, Fonseca C, Baraúna R, Caetano T, *et al*. Pathogenicity of shiga toxin-producing *Escherichia coli* (STEC) from wildlife: should we care? *Sci Total Environ* 2022;812:152324.
30. Montero DA, Canto FD, Velasco J, Colello R, Padola NL, *et al*. Cumulative acquisition of pathogenicity islands has shaped virulence potential and contributed to the emergence of LEE-negative Shiga toxin-producing *Escherichia coli* strains. *Emerg Microbes Infect* 2019;8:486–502.
31. Jenkins C, Perry NT, Godbole G, Gharbia S. Evaluation of chromogenic selective agar (CHROMagar STEC) for the direct detection of shiga toxin-producing *Escherichia coli* from faecal specimens. *J Med Microbiol*2020;69:487–491.
32. Chattaway MA, Dallman TJ, Gentle A, Wright MJ, Long SE, *et al*. Whole genome sequencing for public health surveillance of shiga toxin-producing *Escherichia coli* other than serogroup O157. *Front Microbiol* 2016;7:258.
33. Ashton PM, Perry N, Ellis R, Petrovska L, Wain J, *et al*. Insight into shiga toxin genes encoded by *Escherichia coli* O157 from whole genome sequencing. *PeerJ* 2015;3:e739.
34. Tewolde R, Dallman T, Schaefer U, Sheppard CL, Ashton P, *et al*. MOST: a modified MLST typing tool based on short read sequencing. *PeerJ* 2016;4:e2308.
35. Steyert SR, Sahl JW, Fraser CM, Teel LD, Scheutz F, *et al*. Comparative genomics and stx phage characterization of LEE-negative shiga toxin-producing *Escherichia coli*. *Front Cell Infect Microbiol* 2012;2:133.
36. Dallman T, Ashton P, Schafer U, Jironkin A, Painset A, *et al*. SnapperDB: a database solution for routine sequencing analysis of bacterial isolates. *Bioinformatics* 2018;34:3028–3029.
37. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, *et al*. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015;43:e15.
38. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, *et al*. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020;37:1530–1534.
39. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021;49:W293–W296.
40. UKHSA. Public health operational guidance for Shiga toxin-producing *Escherichia coli* (STEC); 2023. https://www.gov.uk/government/publications/shiga-toxin-producing-escherichia-coli-public-health-management
41. Carroll KJ, Jenkins C, Harvey-Vince L, Mohan K, Balasegaram S. Shiga toxin-producing *Escherichia coli* diagnosed by Stx PCR: assessing the public health risk of non-O157 strains. *Eur J Public Health* 2021;31:576–582.
42. McCarthy SC, Macori G, Duggan G, Burgess CM, Fanning S, *et al*. Prevalence and whole-genome sequence-based analysis of shiga toxin-producing *Escherichia coli* isolates from the recto-anal junction of slaughter-age Irish sheep. *Appl Environ Microbiol* 2021;87:e0138421.
43. Treier A, Stephan R, Stevens MJA, Cernela N, Nüesch-Inderbinen M. High occurrence of shiga toxin-producing *Escherichia coli* in raw meat-based diets for companion animals-a public health issue. *Microorganisms* 2021;9:1556.
44. Miko A, Pries K, Haby S, Steege K, Albrecht N, *et al*. Assessment of shiga toxin-producing *Escherichia coli* isolates from wildlife meat as potential pathogens for humans. *Appl Environ Microbiol* 2009;75:6462–6470.
45. Mora A, López C, Dhabi G, López-Beceiro AM, Fidalgo LE, *et al*. Seropathotypes, phylogroups, Stx subtypes, and intimin types of wildlife-carried, shiga toxin-producing *Escherichia coli* strains with the same characteristics as human-pathogenic isolates. *Appl Environ Microbiol* 2012;78:2578–2585.
46. Blanco J, Blanco M, Blanco JE, Mora A, González EA, *et al*. Verotoxin-producing *Escherichia coli* in Spain: prevalence, serotypes, and

virulence genes of O157:H7 and non-O157 VTEC in ruminants, raw beef products, and humans. *Exp Biol Med*2003;228:345–351.

47. Balière C, Rincé A, Blanco J, Dahbi G, Harel J, *et al*. Prevalence and characterization of shiga toxin-producing and enteropathogenic *Escherichia coli* in shellfish-harvesting areas and their watersheds. *Front Microbiol* 2015;6:1356.
48. Sánchez S, Beristain X, Martínez R, García A, Martín C, *et al*. Subtilase cytotoxin encoding genes are present in human, sheep and deer intimin-negative, shiga toxin-producing *Escherichia coli* O128:H2. *Vet Microbiol* 2012;159:531–535.
49. Son HM, Duc HM, Honjoh KI, Miyamoto T. Identification of the newly identified subtilase cytotoxin-encoding gene (subAB2-2) among clinical shiga toxin-producing *Escherichia coli* isolates. *Can J Microbiol* 2015;61:990–994.
50. Liu Y, Li H, Chen X, Tong P, Zhang Y, *et al*. Characterization of shiga toxin-producing *Escherichia coli* isolated from cattle and sheep in Xinjiang province, China, using whole-genome sequencing. *Transbound Emerg Dis* 2022;69:413–422.
51. Brett KN, Ramachandran V, Hornitzky MA, Bettelheim KA, Walker MJ, *et al*. stx1c Is the most common shiga toxin 1 subtype among shiga toxin-producing *Escherichia coli* isolates from sheep but not among isolates from cattle. *J Clin Microbiol* 2003;41:926–936.
52. Dallman TJ, Greig DR, Gharbia SE, Jenkins C. Phylogenetic structure of shiga toxin-producing *Escherichia coli* O157:H7 from sublineage to SNPs. *Microb Genom* 2021;7:mgen000544.
53. Dallman TJ, Greig DR, Gharbia SE, Jenkins C. Phylogenetic context of shiga toxin-producing *Escherichia coli* serotype O26:H11 in England. *Microb Genom* 2019;7:000551.
54. Butt S, Jenkins C, Godbole G, Byrne L. The epidemiology of shiga toxin-producing *Escherichia coli* serogroup O157 in England, 2009-2019. *Epidemiol Infect* 2022;150:e52.
55. Al Qabili DMA, Aboueisha A-KM, Ibrahim GA, Youssef AI, El-Mahallawy HS. Virulence and antimicrobial-resistance of shiga toxin-producing *E. coli* (STEC) Isolated from edible shellfish and its public health significance. *Arch Microbiol* 2022;204:510.
56. Nguyen TTH, Kikuchi T, Tokunaga T, Iyoda S, Iguchi A. Diversity of the tellurite resistance gene operon in *Escherichia coli*. *Front Microbiol* 2021;12:681175.
57. Taylor DE. Bacterial tellurite resistance. *Trends Microbiol* 1999;7:111–115.
58. Zhang W, Nadirk J, Kossow A, Bielaszewska M, Leopold SR, *et al*. Phylogeny and phenotypes of clinical and environmental shiga toxin-producing *Escherichia coli* O174. *Environ Microbiol* 2014;16:963–976.
59. Commereuc M, Weill F-X, Loukiadis E, Gouali M, Gleizal A, *et al*. Recurrent hemolytic and uremic syndrome induced by *Escherichia coli*. *Medicine* 2016;95:e2050.
60. Friedrich AW, Bielaszewska M, Zhang W-L, Pulz M, Kuczius T, *et al*. *Escherichia coli* harboring shiga toxin 2 gene variants: frequency and association with clinical symptoms. *J Infect Dis* 2002;185:74–84.