Implications of *stx* loss for clinical diagnostics of Shiga toxin-producing *Escherichia coli*

Names of authors:

Thulasika Senthakumaran[1]
Lin Torstensen Brandal[2]
Bjørn-Arne Lindstedt[3]
Silje Bakken Jørgensen[4]
Colin Charnock[5]
Hege Smith Tunsjø[6]

Departments where work was done:

1. Dpt of Multidisciplinary Laboratory Science and Medical Biochemistry, Genetic Unit, Akershus University Hospital, Norway, and Dpt of Life Sciences and Health, Oslo Metropolitan University, Norway
2. Department of Zoonotic, Food- and Waterborne Infections, Norwegian Institute of Public Health, Norway
3. Dpt of Chemistry, Biotechnology and Food Sciences, Norwegian University of Life Sciences, Norway
4. Dpt of Microbiology and Infection control, Akershus University Hospital, Norway
5. Dpt of Life Sciences and Health, Oslo Metropolitan University
6. Dpt of Life Sciences and Health, Oslo Metropolitan University, Norway and Dpt of Microbiology and Infection control, Akershus University Hospital, Norway

Corresponding author:
Hege Smith Tunsjø ORCID 0000-0002-1114-0717
Oslo Metropolitan University
+47 95052752
Email: hetu@oslomet.no
Running title: 'STEC lost shiga toxin'

Abstract

The dynamics related to the loss of *stx* genes from Shiga toxin-producing *Escherichia coli* remain unclear. Current diagnostic procedures have shortcomings in the detection and identification of STEC. This is partly owing to the fact that *stx* genes may be lost during an infection or in the laboratory. The aim of the present study was to provide new insight into *in vivo* and *in vitro stx* loss in order to improve diagnostic procedures. Results from the study support the theory that loss of *stx* is a strain-related phenomenon and not induced by patient factors. It was observed that one strain could lose *stx* both *in vivo* and *in vitro*. Whole genome comparison of *stx*-positive and *stx*-negative isolates from the same patient, revealed that different genomic rearrangements, such as complete or partial loss of the parent prophage, may be factors in the loss of *stx*. Of diagnostic interest, it was shown that patients can be co-infected with different *E. coli* pathotypes. Therefore, identification of *eae*-positive, but *stx*-negative isolates should not be interpreted as "Shiga toxin-lost" *E. coli* without further testing. Growth and recovery of STEC was supported by different selective agar media for different strains, arguing for inclusion of several media in STEC diagnostics.

Keywords

STEC diagnostics, *stx*-loss, EHEC diagnostics, phage excision, selective culture media

Introduction

Shiga toxin-producing *Escherichia coli* (STEC) are foodborne pathogens that can cause a wide range of gastrointestinal and systemic diseases. Shiga toxins are considered to be the major virulence factors of STEC, and are necessary for the development of severe conditions including hemorrhagic colitis and hemolytic uremic syndrome (HUS) [1]. The family of Shiga toxins contains two subgroups, Stx1 and Stx2, and several subtypes, Stx1a, c, d and Stx2a-h, of which Stx2h is the most recently described subtype [2]. Stx2 subtypes vary in their pathogentic potential. Most studies describe Stx2a, c and d as the most virulent [3]. However, severe clinical outcomes have also been described for other subtypes [4,5]. Genes encoding Stx1 and Stx2 are located on different prophages that integrate at specific sites into the *E. coli* genome, a process known as lysogenic conversion. When the phages are silent, their *E. coli* hosts survive as lysogenic strains without expressing the *stx* genes [6]. Environmental stress factors and/or DNA-damaging agents may activate the phage lytic cycle, with concomitant *stx*

expression and release of Stx toxins and phage particles [7,8]. Alternatively, the Stx encoding phage DNA may spontaneously excise from the STEC genome without subsequent cell lysis. The resulting *E. coli*, without integrated *stx*, is termed 'STEC lost shiga toxin' (STEC-LST). STEC and STEC-LST may co-exist *in vivo* and recycle Stx-encoding phages so that they can exist as *stx*-negative variants and then convert back to *stx*-positive forms [9-11]. The Stx-encoding phage may also be lost during laboratory passages [11,12].

In clinical diagnostics, the presence of STEC is usually confirmed by PCR-based detection of *stx* in cultured *E. coli* or directly in stool samples. In the latter instance, positive results are usually followed up by culturing *E. coli* from stools and performing *stx*-PCR on isolates. Excision or loss of the *stx* prophage *in vivo* or during laboratory sample processing, complicates STEC diagnostics. STEC with intact prophages may be low in numbers and, therefore, difficult to detect. Furthermore, induction into the lytic cycle with subsequent cell disruption adds to the complexity of STEC diagnostics. The aim of the present study was to provide new insight into *in vivo* and *in vitro stx*-loss in order to improve diagnostic procedures. Loss of *stx* was investigated in patient stool samples at different stages of infection and after inoculation onto agar, using qPCR. Whole genome sequencing of *stx*-positive and *stx*-negative bacterial isolates was performed to gain insight into genomic rearrangements that had occurred. Finally, different selective agars were tested for recovery of STEC in the presence of STEC-LST and background flora. Based on our findings, we discuss strategies to improve STEC diagnostics.

Materials and Methods

Patient samples and clinical information

The period of investigation was from 2013-2014. Patients included in the study were either hospitalized at Akershus University Hospital, Norway (Ahus) or in primary health care. All patients were suffering from infectious diarrhea, which was in some cases bloody. Diagnostic samples were received at the Ahus Department of Clinical Microbiology and Infection Control. When STEC was detected in a sample, the patient, or if a juvenile the parents of the patient, were invited to participate in the study. Written consent was obtained from all participants. The study was approved by the Data protection manager at Ahus (Project

number12-042) and by the Regional Committees for Medical and Health Research Ethics (REK), South East, Norway (Project number 2012-102).

Initial diagnostic procedures

Stool samples were investigated for the presence of STEC and other gastrointestinal pathogens using a commercial CE-labeled PCR kit (RIDA®GENE EHEC/EPEC real-time PCR, R-Biopharm AG, Darmstadt, Germany) [13]. Samples that tested positive for *stx* were cultured to recover the STEC isolate, which was then verified, serotyped and further characterized at the National Reference Laboratory for Enteropathogenic Bacteria (NRL) at the Norwegian Institute of Public Health. Stx subtype was determined as described by Scheutz *et al*. [14] with minor modifications. Patients with STEC serotypes known to cause severe infection and subtypes *stx2a*, *stx2c*, *stx2d* and *stx1* were included in the study. STEC subtypes *stx2b* and *stx2e-g* were excluded as they have only rarely been associated with clinical infections in Norwegian patients [3]. STEC subtype *stx2h* was not identified in Norway at the time of the study. Two or three follow-up stool samples were collected from each patient at intervals of approximately one week during the period of illness. Stool samples were stored at -80˚ C until processed.

Recovery of STEC on different culture media

Samples included in the study were cultured onto CHROMagar STEC (CHROMagar Microbiology, Paris, FR) and lactose agar to quantify STEC and putative STEC-LST. Lactose agar, an in-house medium for detection of Enterobacteriaceae, contains tryptose agar base, lactose, sodium chloride and bromothymol blue. For each sample, 16-48 different colonies growing on the agars were examined for *stx, eae* and the corresponding O-serogroup (target genes *wzx* or *wzy*), H-serogroup (target gene *fliC*) or lysozyme P (*lysP*) using real time PCR [13,15-19]. Primers and probes are described in Table 2. QuantiFast Pathogen PCR /IC Kit (Qiagen, Hilden, Germany) was used for TaqMan probe assays, while SYBR Premix Ex Taq (Takara Bio, Inc., Otsu, Japan) was used for the SYBR green assays.

Loss of *stx* during *in vitro* culture

To investigate if *stx* was lost during the *in vitro* culture step, the amount of *stx* relative to its corresponding STEC O-serogroup target gene (alternatively H-serogroup/*lysP*) was determined before and after culture on lactose agar. DNA was extracted from stool samples (100 µl) in Cary-Blair transport medium (Copan Italia S.P.A, Brescia, Italy) or RNA*later*™ stabilization reagent (Qiagen, Hilden, Germany) using QIAsymphony (Qiagen) as previously described [13]. Following stool culture (100 µl) on lactose agar, all the colonies on the plate were suspended in 5 ml PBS, and 200 µl of this suspension was used for DNA extraction using the QIAsymphony protocol. qPCR was performed in triplicates (Table 2) and the relative quantity of *stx* (ΔCt *stx* - O-serogroup target gene) in the stool sample versus colonies on lactose agar was determined for every patient. Four patients were excluded from these analyses due to the unsuccessful recovery of STEC (Patients 3 and 9) or insufficient amounts of sample material (Patients 8 and 10).

Loss of *stx* across stages of infection

Follow-up stool samples from each patient were investigated for *stx* and the O-sergroup target gene (alternatively H-serogroup or *lysP*) that was detected in the initial sample. At intervals of approximately one week from the onset of infection the relative quantities of *stx* (ΔCt *stx* - O-serogroup target gene) were determined and compared to the primary diagnostic sample (time point 1). Patients 2, 3, 4, 9 and 12 were not included in these analyses as follow-up samples were not available.

Whole genome sequencing of *stx*-positive and *stx*-negative isolates

When *stx*-positive and *stx*-negative *E. coli* of the same serogroup were identified in a sample, whole genome sequencing (WGS) was performed for comparison of isolates. DNA was extracted with QIAGEN Genomic-tip 100/G (Qiagen) and a library was prepared using the Pacific Biosciences 20 kb library preparation protocol (Pacific Biosciences, Menlo Park, CA). Size selection of the final library was performed using BluePippin (Sage Science, Beverly, MA, USA) with 10 kb cut-off. The library was sequenced using a Pacific Biosciences RS II instrument employing P6-C4 chemistry with 360 minutes movie time. One SMRT cell was used for sequencing.

158 Bioinformatic analysis of whole genome sequence data

159 Reads were assembled using HGAP v3 (Pacific Biosciences, SMRT Analysis Software
160 v2.3.0). Only subreads longer than 8 kb were used for assembly. Minimus2 software of Amos
161 package was used to circularize contigs which were confirmed by dotplot to contain the same
162 sequence at the beginning and end of the contig. RS_Resequencing 1 software **(**SMRT
163 Analysis v2.3.0) was used to map reads back to assembled and circularized sequences. Low-
164 coverage contigs were removed from the genome sequence before the mapping. Fasta files
165 were submitted to the Centre for Genomic Epidemiology (CGE)
166 (http://www.genomicepidemiology.org/) [20] for species identification with PathogenFinder
167 [21], serotype- and multilocus sequence type identification with SerotypeFinder and MLST
168 Finder [22,23] and the detection of plasmids, and virulence genes with PlasmidFinder and
169 VirulenceFinder [24,25] (Table 3). The assembled chromosome sequences of each pair of *stx*-
170 positive and *stx*-negative *E. coli* of the same serogroup were compared using MAUVE
171 genome comparison tool version 2.4.0 [26]. Core genome MLST (cgMLST) was performed in
172 Ridom SeqSphere+ version 5.1.0 (Ridom GmbH, Germany). Briefly, raw sequence reads
173 were trimmed until an average base quality of 30 was reached in a window of 20 bases, and
174 *de novo* assembly was performed using Velvet version 1.1.04 with default settings. The
175 integrated *Escherichia coli* cgMLST scheme v1 from Enterobase
176 (https://enterobase.warwick.ac.uk/species/index/e.coli) was used. The allelic profiles of the
177 isolates were visualised as a minimum spanning tree using the parameter "pairwise ignoring
178 missing values".

179

180

181 **Results**

182 STEC isolates and their recovery on different culture media

183 Thirteen patients were enrolled in the study and 12 *stx*-positive isolates were isolated from 11
184 of these patients. Two different STEC were identified from one of the patients. In the case of
185 two patients no STEC were cultured from the samples. The 12 *stx*-positive isolates
186 represented four different O-serogroups. One isolate was an unknown O-serogroup and one
187 isolate was identified as *Escherichia albertii* [27] (Table 1). CHROMagar STEC was
188 successful in selecting for STEC O157 and suppressed growth of commensal *E. coli*. STEC
189 O157:H7 from patient 13 was recovered in pure culture on CHROMagar STEC, while it was
190 suppressed by commensal flora on lactose agar in the primary cultivation step (Fig. 1).

CHROMagar STEC was also successful in selecting for STEC O103 from two patient samples, although a third STEC O103 was not initially found using this medium. STEC O26:H11 was not able to grow on CHROMagar STEC, while lactose agar supported growth of all STEC in this study (Fig. 1).

Co-existence of STEC, STEC-LST and other *E. coli* pathotypes

In samples from three different patients, a co-existence of *stx2*-positive/*eae*-negative and *stx2*-negative/*eae*-positive *Escherichia* was found. In one of these patients (p5), 14 of 40 colonies from lactose agar were *stx2*- negative *E. coli* O145:H28, while only one of 40 colonies was *stx2*-positive *E. coli* O145:H28. Twenty-five colonies from lactose agar were commensal *E. coli* (Fig. 1). In samples from patient 7, one colony of *stx2*-positive *E. albertii* and four colonies of *stx2*-negative *E. albertii* were identified out of 34 colonies tested from lactose agar (Fig. 1).

Pairs of *stx*-positive and *stx*-negative colonies of the same serogroup were subjected to whole genome sequencing. Bioinformatic analysis identified the pair of O145:H28 isolates as *E. coli* of the same MLST type. Both isolates contained the same plasmids and virulence genes, and differed only in their Stx-encoding genes (Table 3). The pair of *E. albertii* isolates was also shown to contain identical virulence genes, except for *stx* (Table 3). Genome analysis using MAUVE showed that the *stx*-negative O145:H28 isolate was lacking a ~9000 base pair partial sequence region of the *stx*-prophage (Fig. 2a). The *stx*-negative *E. albertii* was missing approximately 62 000 base pairs (corresponding to a complete prophage DNA sequence) present in its *stx*-positive counterpart (Fig. 2b). cgMLST using 2513 targets detected illustrated 0 and 1 allelic differences between the *stx*-negative and *stx*-positive *E. coli* O145:H28 (patient 5) and *stx*-negative and *stx*-positive *E. albertii* (patient 7), respectively (Fig. 3). Whole genome sequences of the isolates have been submitted to the European Nucleotide Archive Study ID PRJEB27634 (ERP109742).

Samples from patient 1 also contained a co-existence of *stx2*-positive/*eae*-positive and *stx*-negative/*eae-positive E. coli*. In these samples, the isolates were found to be of different serotypes. Of 24 colonies tested from lactose agar, six were STEC O26:H11, seven were Enteropathogenic *E. coli* (EPEC) O21, eight were EPEC of unknown serotype and three were Enteoaggregative *E. coli* (EAEC) O104:H4. On CHROMagar STEC, only EPEC O21 and EAEC O104:H4 were detected (Fig. 1).

Loss of *stx* across stages of infection and after *in vitro* culture

The general pattern for all patients was a decrease in both *stx* and *O/H/lysP* from time point 1 to time point 4, and hence a reduction of STEC over time. In two patients (patients 5 and 8), neither *stx* nor the O serogroup target gene were detected in the follow-up samples. In five patients (patients 1, 6, 8, 10 and 11), the relative quantities of *stx* were stable or higher in the follow-up samples (Fig. 4). Samples from patient 7 diverged from this pattern. Here, *stx* was not detected at time points 2, 3 and 4, while *lysP* (*E. albertii* specific target) was detected at every time point, indicating the presence of *stx*-negative *E. albertii* and *in vivo stx* loss (Fig. 4). Subsequent to agar culture, decreased *stx* quantities relative to the O serogroup target gene were observed for three of the samples (patient 1, 5 and 7, log2 fold change >2.5). This suggests *in vitro* loss of *stx* (Fig. 5).

Discussion

The aim of the present study was to provide new insight into *in vivo* and *in vitro stx*-loss in order to improve diagnostic procedures. A proposed theory for loss of Stx-encoding phages from STEC is that this offers a selective advantage for the cell and favors bacterial survival [11]. Both strain-related and patient-related factors, as well as environmental factors may influence this process. Most available information at the present time support the theory that *stx*-loss is related to serotype or *stx* subtype. Mellmann and colleagues [11] documented progressive *stx* loss in seven different patients infected with STEC O26:H11 and O157:NM. Another study from the same group found that 5% of HUS patients had shed STEC-LST O26:H11, O103:H2, O145:H28 and O157:H7 by the time of testing [10]. *stx* loss was identified in STEC O145:H28 also in the present study, and indicated, but not verified in STEC O26:H11. Two different STEC isolates were identified in one of the patient samples. These were STEC O145:H28 containing *stx2a* and STEC O103:H2 containing *stx1*. Only STEC O145:H28 existed as *stx*-positive and *stx*-negative variants, providing support for the hypothesis that *stx*-loss is related to strain or *stx* type, and not induced by patient factors. In our work, *stx* loss was also detected in a patient infected with *stx2a*-positive *E. albertii*. This isolate was obtained from a patient suffering from bloody diarrhea [27]. To our knowledge, this study is the first to illustrate that *stx2* may be lost from *E. albertii*. Several studies have

shown that *stx* genes are more stably maintained in STEC O157:H7 strains than in non-O157 strains [11,28]. This applies also to the STEC O157 isolates in the present study. Loss of *stx1* is less well known, and our data support the idea that *stx* is lost only in certain serotypes and *stx*-subtypes.

The median length of STEC carriage was in a recent study measured to 24 days [29]. The present work detected *stx* after three or four weeks in samples from seven patients, but for all patients, a progressive reduction of STEC was observed. For the patient with *E. albertii*, *stx* was not found seven days after the onset of disease, whereas *lysP* (*E. albertii* specific target) was detected at seven and 12 days. This finding suggests *stx* loss *in vivo*. Following laboratory culture of primary samples on agar plates, the sample with *E. albertii* displayed decreased quantities of *stx* relative to *lysP*, indicating *stx* loss also *in vitro*. Several studies have shown that STEC is prone to loss of *stx* after *in vitro* manipulation [12, 30]. Joris *et al.* [28] showed that *stx* genes may be lost already during the first subcultivation step; the present study illustrates that not only subcultivation, but primary cultivation may have the same effect. We cannot exclude that *in vivo* coexistence of *stx*-positive and *stx*-negative cells may have resulted in bias during the culture stage, as preexisting *stx*-negative variants may be more easily cultured than *stx*-positive variants. A progressive *stx*-loss in this strain in a plausible explanation.

Whole genome sequencing demonstrated that different genomic rearrangements may lead to *stx* loss. Isolates of *stx2*-positive *E. albertii* and *stx2*-negative *E. albertii* were sequenced and bioinformatics analysis showed that these were the same strain, with only one allelic difference using wgMLST. The *stx2*-negative isolate had lost the entire *stx2* encoding phage. In this strain, free bacteriophages may have co-existed with *E. albertii in vivo* and bidirectional conversion of the Stx phage between *stx*-positive and *stx*-negative variants may have occurred [9]. Results from the present study also suggest *in vitro* loss of *stx* in samples from the patient with *stx*-positive and *stx*-negative *E. coli* O145:H28. Whole genome sequencing revealed that these were the same strain, but that one of them was missing some of the *stx* prophage, including the *stx* genes. Remnants of the prophage, including the late gene regulator Q and the genes encoding the phage structural proteins, were still present in the *stx*-negative isolate. This would be in line with a different genomic rearrangement rather than loss of the complete prophage. Since the prophages are not intact when the phage DNA is excised, these are not likely to be transferred to new cells. The reason for only partial loss of the Stx

prophage is not clear. Theoretically, it could be related to bacterial survival and adaptation and lower virulence expression. The patient from which this isolate was identified, suffered from bloody diarrhea and was hospitalized for four days. He was co-infected with STEC O103:H2. Of diagnostic interest, the results illustrate that sequence analysis of more STEC-LST isolates may reveal new diagnostic targets. For example, a PCR-assay for the late gene regulator Q could be used to detect remnants of Stx prophages and identify *stx* loss in some *E. coli* isolates.

Culture on selective agars

STEC diagnostics is often based on detection of *stx* in *E. coli* isolates or directly in stool samples using PCR. An indication of the presence of STEC in a patient sample may be the patient's clinical presentation or a positive *stx* PCR-result from stool samples. A routine procedure for STEC detection used at many hospital laboratories is to test 2-8 colonies isolated from stool samples for the presence of *stx* [13,31].

It can be technically challenging to culture low quantities of STEC in samples with co-occurring STEC-LST and other *E. coli*. In this study, analysis of the two patient samples with pairs of *stx*-positive and *stx*-negative *E. coli* showed that only a minority of the colonies growing on lactose agar and CHROMagar STEC were STEC. The majority of colonies growing from these samples were STEC-LST, competing microflora or commensal *E. coli*. Therefore, detection of 2-8 *stx*-negative colonies in a *stx*-positive stool sample should not lead to termination of the STEC "search". STEC isolates are likely to be present in samples where *stx* loss has occurred, although they will probably exist in small numbers. Furthermore, it would be unwise to assume that *stx*-negative, *eae*-positive colonies are possible STEC-LST without further analysis. In this study, one patient was co-infected with STEC O26:H11, EPEC O21, an EPEC of unknown serotype and EAEC O104:H4. Only EPEC O21 and EAEC O104:H4 grew on CHROMagar STEC. If CHROMagar STEC were to be used as the only culture medium, the *eae*-positive EPEC O21 could be mistaken for STEC-LST. Although several studies have shown that CHROMagar STEC is a suitable medium for STEC O26 [32-34], it is clear that the commonly used selective culture media for STEC do not support growth of all STEC variants. Conversely, the less selective lactose agar also supports growth of commensal *E. coli*. In the present study, commensal *E. coli* suppressed STEC O157:H7 from one sample at the primary cultivation stage. Dual plating on lesser and more selective agars should be performed if STEC isolates are

not recovered after a positive *stx* PCR-result from stool or mixed culture. Our data illustrate that a high number of colonies need to be screened for *stx* if the patient's clinical presentation suggests the presence of STEC.

The present work also showed that STEC may not be recovered if lytic induction occurs during cultivation. In two of the samples (patient 3 and 9), a positive *stx*-PCR was obtained directly from stool and from a culture swipe from lactose agar. Representative samples of the colonies on the plate were tested for *stx*, however, no STEC isolates were identified. It has previously been shown that free Stx phages may exist in patient stool samples and lead to positive *stx*-PCR results [35]. Since *stx* was detected in mixed culture on agar and not only from stool samples in patient 3 and 9, the results are in line with lytic induction upon subculture, rather than the presence of free Stx phages in the sample. In such cases, DNA from culture swipes could be used to search for common STEC serotypes or other genetic STEC markers. Although single isolates cannot be characterized using this approach, a possible STEC infection may be identified and the STEC serotype recorded for infection control purposes. If *stx* is not detected in culture swipes, free Stx phages are more likely to have caused the *stx* positive PCR result.

The present study has some limitations which should be considered. The fecal samples had been frozen prior to analysis, which may have influenced *stx* loss. The small sample size of the study is also a limiting factor.

Acknowledgements

We thank colleagues at the Department of Microbiology and Infection Control at Ahus for laboratory assistance. Parts of the contents have been presented in a report to the National Institute of Public Health.

The sequencing service was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo and supported by the "Functional Genomics" and "Infrastructure" programs of the Research Council of Norway and the Southeastern Regional Health Authorities

Compliance with Ethical Standards

This study received funding from Akershus University Hospital, National Institute of Public Health and Oslo Metropolitan University.

The authors declare no conflict of Interest.

The study was approved by the Data protection manager at Ahus (Project number12-042) and by the Regional Committees for Medical and Health Research Ethics (REK), South East, Norway (Project number 2012-102).

Informed consent was obtained from all participants.

References

1. Tarr PI, Gordon CA, Chandler WL (2005). Shiga-toxin-producing Escherichia coli and haemolytic uremic syndrome. Lancet 365:1073-86

2. Bai X, Fu S, Zhang J, Fan R, Xu Y, Sun H, He X, Xu J, Xiong Y (2018). Identification and pathogenomic analysis of an *Escherichia coli* strain producing a novel Shiga toxin 2 subtype. Scientific Reports 8:6756. DOI:10.1038/s41598-018-25233-x

3. Naseer U, Løbersli I, Hindrum M, Bruvik T, Brandal LT (2017). Virulence factors of Shiga toxin-producing Escherichia coli and the risk of developing haemolytic uremic syndrome in Norway, 1992-2013 Eur J Clin Microbiol Infect Dis 36:1613-20

4. Fasel D, Mellmann A, Cernela N, Hächler H, Fruth A, Khanna N, Egli A, Beckmann C, Hirsch HH, Goldenberger D, Stephan R (2014). Hemolytic Uremic Syndrome in a 65-year-old male linked to a very unusual type of *stx*$_{2e}$ - and *eae*-harboring O51:H49 Shiga toxin-producing *Escherichia coli*. JCM 52(4):1301-3

5. Saupe A, Edel B, Pfister W, Löffler B, Ehricht R, Rödel J (2017). Acute diarrhea due to a Shiga toxin 2e-producing *Escherichia coli* O8:H19. JMM Case Reports 4. DOI:10.1099/jmmcr.0.005099

6. Herold S, Karch H, Schmidt H (2004). Shiga toxin-encoding bacteriophages — genomes in motion. Int J Med Microbiol 294:115–21

7. de Sablet T, Bertin Y, Vareille M, Girardeau JP, Garrivier A, Gobert AP, Martin C (2008). Differential expression of stx2 variants in Shiga toxin-producing Escherichia coli belonging to seropathotypes A and C. Microbiology 154(Pt 1):176-86

8. Waldor MK, Friedman DI (2005). Phage regulatory circuits and virulence gene expression. Curr Opin Microbiol 8(4):459-65

9. Bielaszewska M, Prager R, Köck R, Mellmann A, Zhang W, Tschäpe H, Tarr PI, Karch H (2007). Shiga toxin gene loss and transfer in vitro and in vivo during enterohemorrhagic Escherichia coli O26 infection in humans. Appl Environ Microbiol 73(10):3144-50

10. Bielaszewska M, Köck R, Friedrich AW, von Eiff C, Zimmerhackl LB, Karch H, Mellmann A (2007). Shiga Toxin-Mediated Hemolytic Uremic Syndrome: Time to Change the Diagnostic Paradigm? PLoS ONE 2(10):e1024

11. Mellmann A, Lu S, Karch H, Xu JG, Harmsen D, Schmidt MA, Bielaszewska M (2008). Recycling of Shiga toxin 2 genes in sorbitol-fermenting enterohemorrhagic Escherichia coli O157:NM. Appl Environ Microbiol 74(1):67-72

12. Karch H, Meyer T, Rüssmann J (1992). Frequent loss of Shiga-like toxin genes in clinical isolates of Escherichia coli upon subcultivation. Infect Immun 60:3464-67

13. Tunsjø HS, Kvissel AK, Follin-Arbelet B, Brotnov B-M, Ranheim TE, Leegaard TM (2015). Suitability of stx-PCR directly from fecal samples in clinical diagnostics of STEC. APMIS 123(10):872-8

14. Scheutz F, Teel LD, Beutin L, Piérard D, Buvens G, Karch H, Mellan A, Caprioli A, Tozzoli R, Morabito S, *et al.* (2012). Multicenter Evaluation of a Sequence-Based Protocol for Subtyping Shiga Toxins and Standardizing Stx Nomenclature. JCM 50(9):2951-63

15. Nielsen EM, Andersen MT (2003). Detection and characterization of verocytotoxin-producing Escherichia coli by automated 50 nuclease PCR assay. J Clin Microbiol 41:2884–93

16. Lin A, Sutton O, Lau HK, Wong E, Hartman G, Lauzon CR (2011). O-serogroup specific real-time PCR assays for the detection and identification of nine clinically relevant non O-157 STECs. Food Microbiology 28:478-483

17. Guy RA, Tremblay D, Beausoleil L, Harel J, Champagne MJ (2014). Quantification of E. coli O157 and STEC in feces of farm animals using direct multiplex real time PCR (qPCR)

427 and a modified most probable number assay comprised of immunomagnetic bead separation
428 and qPCR detection. J Microbiol Methods 99:44-53
429
430 18. Beutin L, Delannoy S, Fach P (2015). Sequence Variations in the Flagellar Antigen Genes
431 fliCH25 and fliCH28 of Escherichia coli and Their Use in Identification and Characterization
432 of Enterohemorrhagic E. coli (EHEC) O145:H25 and O145:H28. PLoS One 10(5):e0126749
433
434 19. Hyma KE, Lacher DW, Nelson AM, Bumbaugh AC, Janda JM, Strockbine NA, Young
435 VB, Whittam TS (2005). Evolutionary Genetics of a New Pathogenic *Escherichia* Species:
436 *Escherichia albertii* and Related *Shigella boydii* Strains. J Bacteriol 2005;187(2): 619–628.

437 20. Larsen MV, Cosentino S, Lukjancenko O, Saputra D, Rasmussen S, Hasman H, Sicheritz-
438 Pontén T, Aarestrup FM, Ussery DW, Lund O (2014). Benchmarking of methods for genomic
439 taxonomy. J Clin Microbiol 52(5):1529-39.

440 21. Cosentino S, Voldby Larsen M, Møller Aarestrup F, Lund O (2013). PathogenFinder –
441 Distinguishing Friend from Foe using Bacterial Whole Genome Sequencing Data. PlosOne
442 8(10):e77302.
443
444 22. Joensen KGAM, Tetzschner A, Iguchi FM, Aarestrup M, Scheutz F (2015). Rapid and
445 easy in silico serotyping of Escherichia coli using whole genome sequencing (WGS) data J
446 Clin Microbiol 53(8):2410-26.

447 23. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L,
448 Sicheritz-Pontén T, Ussery DW, Aarestrup M, Lund O (2012). Multilocus Sequence Typing
449 of Total Genome Sequenced Bacteria. J Clin Micobiol 50(4):1355-61.

450 24. Carattoli A, Zankari E, Garcia-Fernandez A, Voldby Larsen M, Lund O, Villa L,
451 Aarestrup FM, Hasman H (2014). In Silico Detection and Typing of Plasmids using
452 PlasmidFinder and Plasmid Multilocus Sequence Typing. Antimicrob Agents and Chemother
453 58(7), 3895-3903.
454
455 25. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM (2014).
456 Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection
457 of verotoxigenic Escherichia coli. J Clin Microbiol 52(5):1501-10.

26. Darling AC, Mau B, Blattner FR, Perna NT (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403.

27. Brandal LT, Tunsjø HS, Ranheim TE, Løbersli I, Lange H, Wester AL (2015). Shiga toxin 2a in *Escherichia albertii*. JCM 53(4):1454-55)

28. Joris M-A, Verstraete K, De Reu K, De Zutter L (2011). Loss of *vtx* Genes after the First Subcultivation Step of Verocytotoxigenic *Escherichia coli* O157 and Non-O157 during Isolation from Naturally Contaminated Fecal Samples. Toxins (Basel) 3(6): 672–77.

29. Matussek A, Jernberg C, Einemo I-M, Monecke S, Ehricht R, Engelmann I, Löfgren S, Mernelius S (2017). Genetic make-up of Shiga toxin-producing Escherichia coli in relation to clinical symptoms and duration of shedding: a microarray analysis of isolates from Swedish children. Eur J Clin Microbiol Infect Dis 36:1433-41

30. Shaikh N, Tarr PI (2003). Escherichia coli O157:H7 Shiga toxin-encoding bacteriophages: integrations, excisions, truncations, and evolution-ary implications. J Bacteriol 185:3596–3605.

31. Strategy report of enteropathogens, Norwegian public health, 2011. Available from https://www.fhi.no/globalassets/dokumenterfiler/rapporter/strategirapporter/strategirapport-nr-10-1996-bakteriologiske-faecesundersokelser-med-revisjoner-fra-2007-og-2011-.pdf

32. Hirvonen JJ, Siitonen A, Kaukoranta SS (2012). Usability and performance of CHROMagar STEC medium in detection of Shiga toxin-producing Escherichia coli strains. J Clin Microbiol 50(11):3586-90.

33. Gill A, Huszczynski G, Gauthier M, Blais B (2014). Evaluation of eight agar media for the isolation of shiga toxin-producing Escherichia coli. J Microbiol Methods 96:6-11.

34. Gouali M, Ruckly C, Carle I, Lejay-Collin M, Weill F-X (2013). Evaluation of CHROMagar STEC and STEC O104 Chromogenic Agar Media for Detection of Shiga Toxin-Producing Escherichia coli in Stool Specimens. J Clin Microbiol 51(3):894-900.

35. Martinez-Castillo A, Quirós P, Navarro F, Miró E, Muniesa M (2013). Shiga toxin 2-encoding bacteriophages in human fecal samples from healthy individuals. Applied Environ Microbiol 79:4862-68.

.

Tables and Figures

Table 1. Disease characteristics, subtyping and serotyping results for each patient

Patient ID	Age	Sex	Diarrhea	Duration of symptoms (days)	Bloody diarrhea	Fever	Leukocytes max	Thrombocytes min.	Haemoglobin min.	Serum creatinine max.	Antibiotic treatment	Duration of hospitalization (days)	Duration symptoms before hospitalization (days)	Subtyping result stx /eae	Number of follow-up samples	Serotyping result
1	<1	M	yes	2	yes	no	5.7	176	8.9	31	no	2	4	*stx2a, eae*	3	O26:H11
2	8	K	yes	-	no	no	13.4	363	13.1	-	no	1	5	*Stx1*	0	O103:
3	1	M	yes	1	no		28.6	326	13.8	25	no	1	1	*Stx1*	0	O?H?
4	5	K	yes	-	yes	yes	9.6	286	12.9	32	no	1	2	*Stx1*	0	O103:
5	47	M	yes	4	yes	no	10.1	272	14.7	83	yes	4	2	*stx1/eae* *stx2a, eae*	2	O103:H2/ O145:H28
6	63	M	yes	>10	yes		3.3	83	7.5	140	no	23	>10	*stx2a, eae*	3	H25
7	58	M	yes	7	yes	no	9.4	302	16.6	73	no	1	7	*stx2a, eae*	3	*E. albertii*
8	19	F	yes	3	no	no	19.1	387	13.6	60	no	1	3	*stx1/2c, eae*	3	O157:H-
9	10	F	yes	14	yes	no								*Stx1*	0	O?:H?
10	21	F	yes	4	yes	yes	12.3	322	14.1	84	no	0	-	*stx1,eae*	3	O103:H2
11	42	F	yes	7	no	no					no	0	-	*stx1, eae*	3	O103:H2
12	70	F	yes	3	yes	no	11.3	198	11.7	54	yes	1	2	*stx1/2c, eae*	0	O157:H-
13	63	F	yes	10	yes	yes	17.4	238	10.1	60	no	0	-	*stx1/2c, eae*	3	O157:H-

513

514 Table 2. PCR Primers and probes used in this study

PCR assay	Target gene	Primer/probe sequence 5' to 3'	Amplicon size (bp)	Ta*	PCR-efficiency	Reference
Stx 1	*stx1*	Fwd: GGATAATTTGTTTGCAGTTGATGTC Rev: CAAATCCTGTCACATATAAATTATTTCGT 6-FAM: CCGTAGATTATTAAACCGCCCTTCCTCTGGA	107	60°C	92	[15]
Stx 2	*stx2*	Fwd: GGGCAGTTATTTTGCTGTGGA Rev: GAAAGTATTTGTTGCCGTATTAACGA Y.yellow: ATGTCTATCAGGCGCGTTTTGACCATCTT	131	60°C	99	[15].
O26	*wzx*	Fwd: TTTTATCTGGCGTGCTATCG Rev: CGGGGTTGCTATAGACTGAA 6-FAM: TGGCACTCT/ZEN/TGCTTCGCCTG	247	52°C	94	[16]
O103	*wzx*	Fwd: GGGCTTGTATTGTACCG Rev: AGTGGCAAACAGCCAACTAC 6-FAM: TCGGGGATT/ZEN/TTCTGCGGATT	169	52°C	97	[16]
O145	*wzy*	Fwd: TGTTCCTGTCTGTTGCTTCA Rev: ATCGCTGAATAAGCACCACT 6-FAM: TGGGCTGCC/ZEN/ACTGATGGGAT	291	52°C	96	[16]
O157	*per*	Fwd: GTACAAGTCCACAAGGAAAG Rev: CTTGTTTCGATGAGTTTATCTGCA 6-FAM: AGGACCGCAGAGGAAAGAGAGGAATT	125	52°C	98	[17]
H25	*fliC*	Fwd: CACAACATYCTTGATAAAGATGG Rev: AACAGAAGCAGCATAGAAGTC 6-FAM: GCAACAGCTGATTATGTTGTTCAGTCAGG	81	60°C	101	[18]
Lys P	*lysP*	Fwd: GGGCGCTGCTTTCATATATTCTT Rev: TCCAGATCCAACCGGGAGTATCAGGA	252	52°C	93	[19]

515 ***Ta: Annealing temperature.**

516

Table 3. Whole genome sequencing of pairs of *stx*-positive and *stx*-negative *Escherichia* spp.

	***stx*-negative *E. coli* O145:H28**	***stx*-positive *E. coli* O145:H28**	***stx*-negative *E. albertii***	***stx*-positive *E. albertii***
Number of reads	94 292	101 347	72 997	63 006
Read length	15 445	16 738	16 527 bp	17 540
Average coverage	182,3	175,6	210,5	72,26
Number of contigs	5	5	1	8
Number of contigs after circularization	4	3	1	1
Contig lengths	5 457 886 bp 115 737 bp 114 722 bp 56 994 bp	5 461 692 bp 109 862 bp 57 341 bp	4 539 208 bp	4 599 602 bp
Concensus accuracy	0.99982	0.99989	0.99979	0.99964
Stx integration site	yecE	yecE	-	wrbA
Serotype[1] MLST type[2,5]	O145:H28 ST-6130	O145:H28 ST-6130	No H/O No ST	No H/O No S
Plasmids[3]	IncFIB(AP001918) IncI2	IncFIB(AP001918) IncI2	none	none
Virulence genes[4]	iha, tccP, nleC, espJ, cif, nleB, efa1, tir, eae, espA, espB, gad, nleA, nleB, nleC, astA, etpD, ehxA	iha, tccP, nleC, espJ, cif, nleB, efa1, tir, eae, espA, espB, gad, nleA, nleB, nleC, astA, etpD, ehxA, STX2A, STX2B	nleB, , cif, espJ, espF, espA, eae, tir, gad, cdtB, gad	nleB, , cif, espJ, espF, espA, eae, tir, gad, cdtB, gad, STX2A, STX2B

1 = SerotypeFinder [22], 2 = MLST Finder [23], 3 = PlasmidFinder [24], 4 = VirulenceFinder [25], 5=cgMLST Ridom SeqSphere+ .

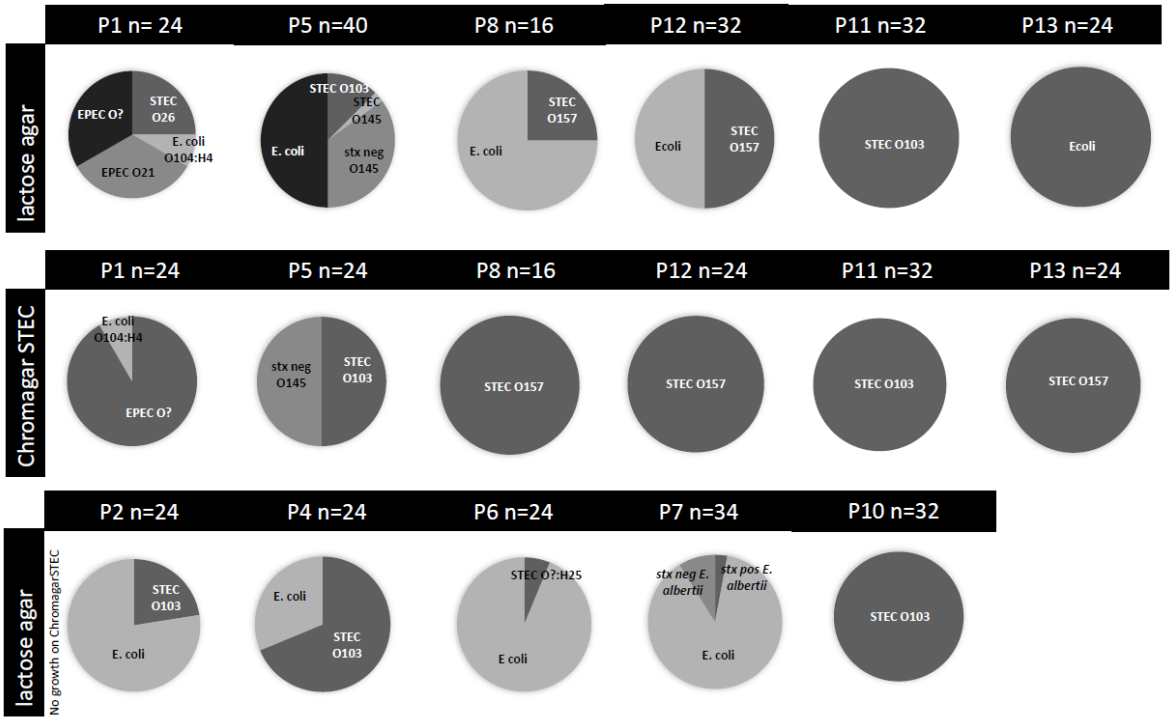
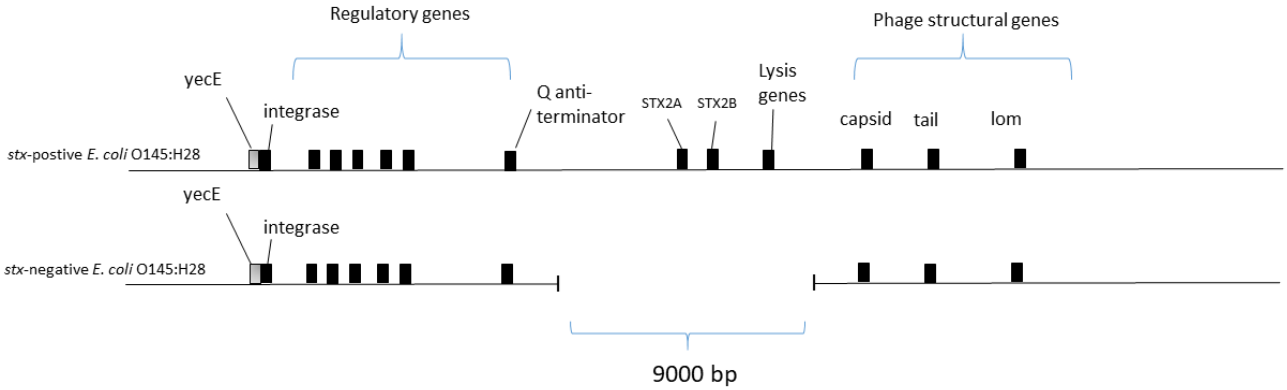


Fig. 1. Recovery of STEC on CHROMagar STEC and lactose agar. In particular, STEC O157 and STEC O103 were recovered on CHROMagar STEC. Lactose agar was necessary for growth of STEC O26:H11 and several other serotypes.

2a



2b

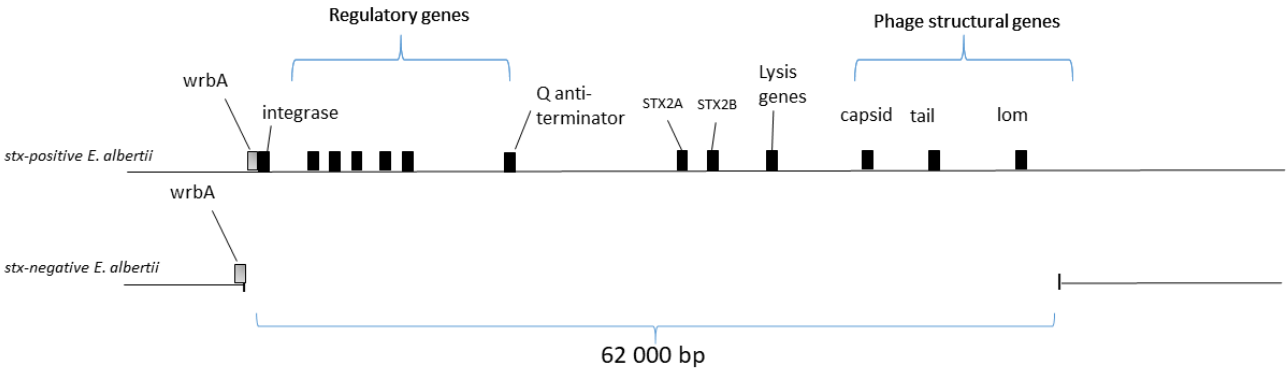


Fig. 2. (a) Graphical presentation of the integrated Stx encoding bacteriophage sequence in *stx*-positive and *stx*-negative isolates of the same serotype. The *stx*-negative *E. coli* O145:H28 isolate is missing a stretch of approximately 9000 basepairs of the Stx encoding bacteriophage compared to the *stx*-positive *E. coli* O145:H28. (b) The *stx*-negative E. *albertii* is missing the complete Stx prohage DNA sequence compared to its *stx*-positive counterpart (62 000 bp).

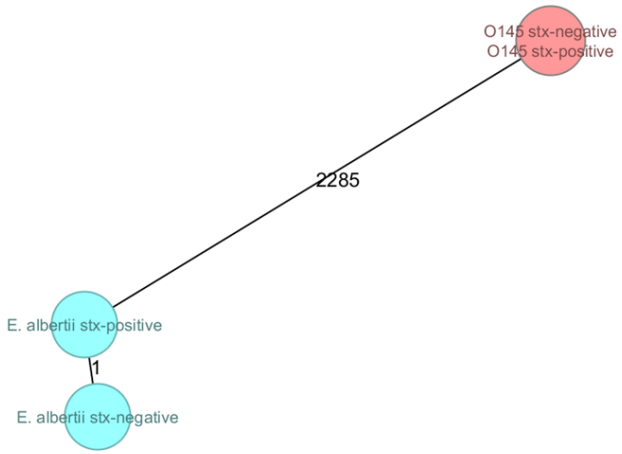


Fig. 3. Minimum-spanning tree based on cgMLST allelci profiles of two *E. coli* and two *E. albertii* isolates (*stx*-positive and *stx*-negative pairs). Each circle represents an allelic profile based on sequence analysis of 2513 targets. The numbers of the connecting lines illustrate the numbers of target genes with differing alleles. Both *E. coli* isolates had 98,4% good cgMLST targets, whereas the *E. albertii* isolates had 92,8 and 94,6% good cgMLST targets, respectively.

Only, one core genome gene showed allelic difference between the *E. albertii* isolates, whereas none differences were observed between the *E. coli* isolates.

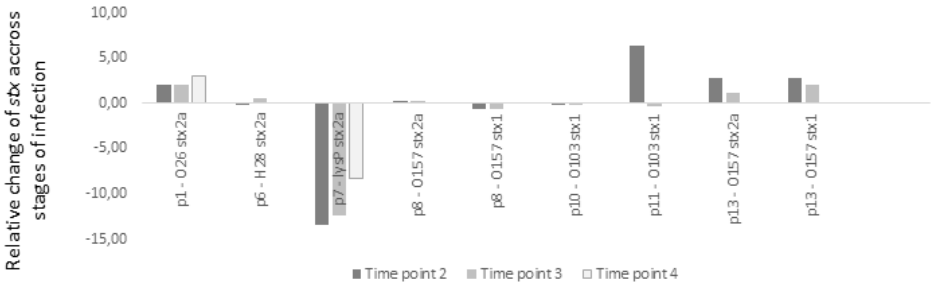


Fig. 4. Relative change (log2 fold change) of *stx* in follow-up samples compared to time point 1 for each patient. Samples from patient 7 were *stx*-negative at time point 2, 3 and 4, while *lysP* was detected at every time point, indicating the presence of *stx*-negative *E. albertii* and loss of *stx in vivo*.

Fig. 5. Relative change (log2 fold change) of *stx* in bacterial growth on agar compared to stool sample from the same patient. Lower quantities of *stx* relative to the O serogroup target gene (or *lysP*) were observed for patient 1, 5 and 7, indicating *stx*-loss *in vitro*.